UNIVERSITÀ DEGLI STUDI DI PADOVA

SEDE AMMINISTRATIVA: UNIVERSITÁ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE DEL FARMACO

SCUOLA DI DOTTORATO IN SCIENZE MOLECOLARI
INDIRIZZO: SCIENZE FARMACEUTICHE
XXV CICLO

Two perspectives to consider Nucleic Acids

Direttore della Scuola: Ch.mo Prof. Antonino Polimeno

Coordinatore d'indirizzo: Ch.mo Prof. Alessandro Dolmella

Supervisore: Ch.ma Prof.ssa Barbara Gatto

Dottoranda: Alice Sosic

31 GENNAIO 2013

A tre persone importanti
Alla mia mamma, la mia forza di sempre
A Barbara, la mia guida ed esempio
A Filippo, il mio futuro

"È proprio quando credete di sapere qualcosa
che dovete guardarla da un'altra prospettiva"
da "L'attimo fuggente"

ABSTRACT

Since their discovery, nucleic acids have been the object of intense and thorough explorations, leading to the understanding of their structure and functions. Their role as genetic information carriers is well known but there are evidences that they are also involved in a series of other less known processes. No longer have nucleic acids to be considered passive structures, but they are dynamic and active macromolecules, able to assume a number of three-dimensional conformations. Precisely, the conformation of nucleic acids, encoded in their primary sequence, is the component that determines the specific and selective interaction with a biological target.

This work is essentially divided in two parts reflecting the dual nature of nucleic acids and aimed at developing two different projects, one where nucleic acids represent a valid target for new potential therapeutic agents and the other where nucleic acids, playing an active role in the protein recognition, are developed as diagnostic agents.

The first part of my PhD research activity is included in a wider project aimed to identify potential antiviral agents able to interfere with the NC chaperone activity. NC has multiple functions during the HIV-1 replication cycle and, in particular, the chaperone activity of NC is critical in reverse transcription. The recombinant full-length NC protein and its biological activity were deeply characterized: the analysis of the NC structure was completed by circular dichroic spectroscopy (CD). Then, the chaperone activity of NC *i.e.* its ability to destabilize the secondary structures of the viral nucleic acids TAR and cTAR was evaluated by a Fluorescence Assay and a FRET assay; NC aggregating activity was instead monitored through gel electrophoresis that allowed us to analyze the TAR/cTAR annealing reaction in the presence of the full-length NC protein.

The effect of compounds on the chaperone activity of NC was investigated: to achieve this goal, a simple, fast and highly reproducible FRET-based assay for the High Throughput Screening (HTS) of inhibitors of the NC nucleic acids destabilization activity was optimized. A wide range of small molecules with different chemical structures from our in-house chemical library were analyzed, and in this work we focus on molecules with intercalating ability.

2,6-dipeptidyl anthraquinones, known DNA and RNA binders, were analysed as potential NC inhibitors: they could interfere with NC chaperone activity thanks to their attitude of stabilize the secondary structures of RNA and DNA sequences involved in the retrotranscription process. The effect of anthraquinones (series Z, GSF and G) on the

chaperone activity of NC was investigated by HTS and the IC_{50} values on TAR and cTAR were determined for each compound. All tested anthraquinones are active both toward TAR and cTAR: the proposed mechanism of action for anthraquinones was therefore studied in more details by Melting Assays (FRET probes coupled to thermal melting experiments). We analyzed the ability of each compound to bind and stabilize the secondary structure of cTAR, TAR and of the TAR/cTAR hybrid. Melting Assay data supported HTS results, since the most active inhibitors of the chaperone NC activity resulted to be also strong TAR and cTAR stabilizers. Finally, we analyzed by native gel electrophoresis the effect of the most promising anthraquinones on the NC-mediated annealing of TAR with cTAR, to verify the effectiveness of the inhibition of the NC chaperone activity during the step subsequent to TAR and cTAR melting. All the tested anthraquinones resulted to be inhibitors of the TAR/cTAR hybrid formation, according with the NC inhibitory activity analyzed by HTS. Hence, viral nucleic acids involved in reverse transcription represent a new validated target for anti-HIV drugs design and 2,6-dipeptydil anthraquinones were identified as highly active anti-NC compounds which could be further analyzed to be developed as new potential antiviral agents.

Moreover, a promising class of anti-HIV agents is represented by compounds which interfere with the Tat-mediated transcription process of the viral replicative cycle. Two classes of compounds were studied as inhibitors of the Tat-TAR complex formation. Quinolones are a promising class of anti-HIV compounds which interfere with the Tat mediated transcription process. To provide new insights into the SAR of antiviral quinolones, WP and HP series were analyzed by Fluorescence Quanching Assay (FQA), a FRET-based platform of screening that allowed the determination of the inhibition constant Ki for each compound. All the tested 6-DFQs of WP series are active as Tat-TAR complex formation inhibitors, with Ki values lower than the control WM5. WP series differs from the control for modifications introduced at the N1 position: these modifications do not result in loss of activity, but rather seem to improve the activity of WP quinolones as inhibitors of the Tat-TAR complex formation. The second class of compounds tested as inhibitors of the Tat-TAR complex formation was constituted by ellagic acid and some synthetic bi- and mono-lactones derivatives. ELs analysis was not possible to be performed by FQA due to a quenching effect observing after the addition of ELs to fluoresceinated Tat. The analysis of the ELs interference on the Tat-TAR complex was therefore performed by EMSA. Compound 1186 can act as Tat-TAR complex formation inhibitor and therefore supports the potential anti-HIV activity of ELs.

The second part of my research work is focused on the development, optimization and characterization of a novel sensing system with therapeutic and diagnostic interest employing nucleic acids aptamers for the detection of proteins involved in diverse diseases. Biosensors based on the use of aptamers for specific recognition of an analyte are called "aptasensors". We developed the Sandwich Aptamer Microarray (SAM) with the aim to replace the traditional immunochemical systems used as diagnostics. SAM is a system ELISA-like, using a similar architecture, but totally based on aptamer technology with a "primary aptamer" as selection element and a "secondary aptamer" as a signal transducer. Two DNA aptamers with distinct protein recognition patterns are used in tandem for the simultaneous recognition of the target protein.

We developed an aptamer-based microarray for human thrombin detection exploiting two non-overlapping DNA thrombin binding aptamers recognizing different exosites of the target protein. The 15-mer aptamer (TBA1) binds the fibrinogen-binding site, whereas the 29-mer aptamer (TBA2) binds the heparin binding domain. Appropriate chemical modifications were introduced in primary and secondary aptamers to adapt them to the format design and to the detection technology. Extensive analysis on the complex formation between human thrombin and modified aptamers was performed by EMSA, in order to verify in solution whether the chemical modifications introduced would affect aptamers/protein recognition. The validated system was then applied to the aptamer microarray, using the solid phase system devised by the solution studies. Our SAM system based on aptamer recognition of an analyte resulted to be efficient and specific, and different detection aptamers and methods can be utilized. The use of the indirect method (TBA2- biotin plus Cy3-streptavidin) allows lower LOD and LOQ when compared to TBA2-Cy5 (direct method) even in the presence of biological fluids. SAM exhibits a limit of detection comparable to other systems described in the literature employing the same aptamers and different technologies but has the advantage of a simple set up.

In view of future applications of SAM in multiplex microarray systems, we enlarged the SAM system analysis on the $VEGF_{165}$ detection using two anti-$VEGF_{165}$ aptamers that recognize different sites of the protein. Following the same approach adopted for thrombin, a deep analysis and optimization of binary and ternary complexes formation between the modified anti-$VEGF_{165}$ aptamers and $VEGF_{165}$ in solution was performed prior to verify the sandwich aptamer formation in solid phase. The sandwich aptamer formation was verified also for $VEGF_{165}$. This pioneering work suggests that SAM, developed for thrombin and $VEGF_{165}$ detection, could represent a useful tool to monitor protein concentration in blood or plasma and represents a potential application of aptamers in diagnostic.

RIASSUNTO

Dal giorno della loro scoperta, più di un secolo fa, gli acidi nucleici sono stati analizzati e studiati approfonditamente in tutti i loro aspetti, cercando di svelare gli ancora numerosi segreti nascosti nella loro struttura per comprenderne appieno le funzioni. Il ruolo più comunemente associato ad essi è sicuramente quello di trasportatori dell'informazione genetica ma numerose evidenze sperimentali ne confermano il coinvolgimento in una serie di altri processi meno noti. Gli acidi nucleici non devono perciò essere concepiti come strutture rigide e passive, ma come molecole dinamiche e attive, in grado di assumere numerose strutture tridimensionali. È proprio la conformazione spaziale che assumono gli acidi nucleici, codificata nella loro sequenza nucleotidica, che determina l'interazione specifica e selettiva con un bersaglio molecolare.

Questo lavoro è essenzialmente diviso in due parti ed è finalizzato alla comprensione della duplice natura degli acidi nucleici, sviluppando due aspetti differenti ma complementari legati ad essi. Essendo correlati a numerosi eventi cellulari, gli acidi nucleici rappresentano un ottimo target per lo sviluppo di nuovi farmaci e allo stesso tempo, svolgendo un ruolo attivo nel riconoscimento di elementi proteici, gli acidi nucleici possono essere sviluppati come agenti diagnostici.

La prima parte del mio lavoro di dottorato fa parte di un ampio progetto volto a identificare potenziali agenti antivirali, la cui azione può essere ricondotta all'inibizione dell'attività di NC mediante l'interazione diretta con il DNA e l'RNA virali. NC svolge numerose funzioni durante il ciclo replicativo di HIV-1 ed in particolare, grazie all'attività di *chaperone*, promuove varie reazioni di *annealing* nel processo di retrotrascrizione. Piccole molecole che siano in grado di riconoscere gli acidi nucleici virali coinvolti in questi passaggi possono essere considerati potenziali inibitori del ciclo replicativo di HIV-1, poiché in grado di interferire con l'attività di NC. La proteina ricombinante NC, full-length, e la sua attività biologica sono state caratterizzate: l'analisi della struttura di NC è stata completata mediante dicroismo circolare (CD). L'attività chaperonica di NC cioè la sua capacità di destabilizzare le strutture secondarie di TAR e cTAR è stata valutata utilizzando un saggio di fluorescenza e un saggio FRET; l'attività aggregante di NC è stata invece monitorata attraverso gel elettroforesi e ci ha permesso di analizzare la reazione di formazione dell'ibrido TAR/cTAR in presenza della proteina NC. È stata quindi studiata la capacità di composti di interferire sull'attività chaperonica di NC: è stato ottimizzato un saggio FRET che è risultato semplice, veloce e altamente riproducibile per l'*High Throughput Screening* (HTS) di inibitori di NC. È stato

possibile analizzare più di cento composti appartenenti a famiglie chimiche molto diverse, ma in questo lavoro ci siamo concentrati su piccole molecole con attività intercalante.

Gli antrachinoni 2,6-dipeptidil sostituiti, noti leganti di DNA e RNA, sono stati analizzati come potenziali inibitori di NC: essi potrebbero interferire con l'attività chaperonica di NC stabilizzando le strutture secondarie di DNA e RNA coinvolte nel processo di retrotrascrizione. Gli antrachinoni delle serie Z, GSF e G sono stati analizzati mediante HTS e i valori di IC_{50} nei confronti di TAR e cTAR sono stati determinati per ciascun composto. Tutti gli antrachinoni testati sono attivi sia su TAR che su cTAR: il meccanismo d'azione proposto per questi composti è stato quindi approfondito mediante saggi di Melting (sonde FRET accoppiate ad esperimenti di melting). Abbiamo analizzato la capacità di ciascun composto di legare e stabilizzare le strutture secondarie rispettivamente di TAR, cTAR e dell'ibrido TAR/cTAR. I risultati ottenuti mediante saggi di Melting supportano i risultati ottenuti mediante HTS, poiché i composti maggiormente attivi come inibitori di NC sono risultati anche dei potenti stabilizzanti di TAR e cTAR. Infine, abbiamo analizzato, mediante gel elettroforesi, l'effetto degli antrachinoni più promettenti sulla reazione di formazione dell'ibrido mediata da NC, per confermare la correlazione tra inibizione dell'attività chaperonica di NC e l'immediatamente successiva formazione dell'ibrido. Tutti gli antrachinoni testati sono risultati inibitori della formazione dell'ibrido TAR/cTAR, in accordo con i risultati di HTS. Quindi, acidi nucleici virali coinvolti nella retrotrascrizione rappresentano un nuovo e validato bersaglio per lo sviluppo di farmaci anti-HIV e gli antrachinoni 2,6-dipeptidil sostituiti sono stati identificati come composti attivi anti-NC che potrebbero essere sviluppati come potenziali agenti antivirali.

Inoltre, una promettente classe di potenziali anti-HIV è rappresentata da composti che interferiscono con il processo di trascrizione, Tat-mediato, del ciclo replicativo virale. Due classi di composti sono state studiate come inibitori della formazione del complesso Tat-TAR. I chinoloni sono una nota classe di farmaci anti-HIV che interferiscono con il processo di trascrizione mediata da Tat. Per fornire nuove informazioni SAR (Structure-Activity Relationship) sui chinoloni antivirali, le serie WP e HP sono state analizzate mediante test di Fluorescence Quanching (FQA), una piattaforma di screening basata su saggi FRET che ha permesso la determinazione della costante di inibizione Ki per ciascun composto. Tutti i 6-DFQs della serie WP testati sono attivi come inibitori della formazione del complesso Tat-TAR, con valori di Ki inferiori rispetto al controllo WM5. La serie WP differenzia dal controllo per modifiche introdotte nella posizione N1: tali modifiche sembrano migliorare l'attività dei chinoloni WP come inibitori Tat-TAR. La seconda classe di composti testati

come inibitori Tat-TAR è costituita da acido ellagico e alcuni derivati sintetici bi-e mono-lattoni (ELs). Non è stato possibile analizzare gli ELs mediante FQA, e l'analisi è stata quindi eseguita mediante EMSA. Il composto 1186 è un buon inibitore della formazione del complesso Tat-TAR e questo risultato perciò supporta l'ipotesi di sviluppare gli ellagici come potenziali antivirali.

La seconda parte del mio lavoro di ricerca è focalizzata sull'ottimizzazione, sviluppo e caratterizzazione di un nuovo sistema sensoristico per la rilevazione di proteine coinvolte in malattie, con interesse diagnostico e terapeutico. Biosensori basati sull'uso di aptameri per il riconoscimento specifico di un analita sono chiamati "aptasensori". Abbiamo sviluppato un Sandwich Aptamer Microarray (SAM) con lo scopo di sostituire i tradizionali sistemi immunochimici utilizzati in diagnostica. SAM è un sistema simil-ELISA, che usa un'architettura simile, ma è totalmente basato sulla tecnologia degli aptameri: un "aptamero primario" è utilizzato come elemento di selezione e un "aptamero secondario" come trasduttore del segnale. Aptameri di DNA che riconoscono differenti esositi della proteina target vengono utilizzati in tandem per il riconoscimento simultaneo della proteina bersaglio.
Il primo target proteico è costituito dalla trombina. Sono noti e ben caratterizzati due aptameri di DNA che legano la trombina in esositi diversi: TBA1 lega il sito di riconoscimento del fibrinogeno, mentre TBA2 lega la trombina a livello del sito di riconoscimento dell'eparina. Gli aptameri sono stati opportunamente modificati per l'ancoraggio alla superficie solida e per la rilevazione del segnale di fluorescenza. Un'analisi estensiva in soluzione della formazione dei complessi binari e ternari aptameri-proteina è stata effettuata mediante EMSA, per verificare se in soluzione le modifiche chimiche introdotte compromettono il riconoscimento della proteina da parte degli aptameri. Poi, è stata verificata la formazione del sandwich in fase solida. SAM è risultato efficace e specifico, e diversi aptameri secondari e diversi metodi di rilevazione possono essere utilizzati. L'utilizzo del metodo indiretto (TBA2-biotina più streptavidina-Cy3) permette di raggiungere LOD e LOQ inferiori rispetto a TBA2-Cy5 (metodo diretto) anche in presenza di fluidi biologici.

Il secondo target proteico che è stato utilizzato è il $VEGF_{165}$: due aptameri di DNA anti-VEGF sono stati recentemente selezionati. Seguendo lo stesso approccio adottato per la trombina, la coppia di aptameri modificati che assicura la formazione a sandwich è stata identificata in soluzione e il sistema è stato poi applicato in fase solida. I risultati ottenuti sono positivi e coerenti con i risultati ottenuti dall'analisi in soluzione. SAM, sviluppato per la detection della trombina e del VEGF, rappresenta un potenziale mezzo per monitorare i livelli

di proteina nel plasma o nel sangue e dimostra quindi la potenzialità dell'uso degli aptameri in campo clinico e diagnostico.

SUMMARY

1. IDENTIFICATION OF POTENTIAL ANTI-HIV AGENTS

2. APTAMER BASED-MICROARRAY FOR PROTEIN DETECTION: A POTENTIAL TOOL FOR DIAGNOSTIC APPLICATIONS

1.

IDENTIFICATION OF POTENTIAL ANTI-HIV AGENTS

1.1 INTRODUCTION

1.1.1 Human Immunodeficiency Virus

The Human Immunodeficiency Virus (HIV) is the causative agent of the Acquired Immune Deficiency Syndrome (AIDS), which, if untreated, is a fatal disease.

HIV infection typically occurs over the course of several years. HIV infects vital cells of the human immune system such as helper T cells (specifically $CD4^+$ T cells), macrophages and dendritic cells. HIV infection is characterized by a gradual decrease in T cell levels (in particular T cells carrying the CD4 marker on their surface which include T helper cells) through three main mechanisms: direct viral killing of infected cells; increased rates of apoptosis in infected cells and killing of infected $CD4^+$ T cells by CD8 cytotoxic lymphocytes that recognize infected cells. When $CD4^+$ T cell number decline below a critical level, cell mediated immunity is lost.

AIDS is the final stage of HIV infection and is used to describe a highly decimated $CD4^+$ T cell count and the appearance of various opportunistic infections as a result of the progressive failure of the immune system. Any secondary condition, symptom, or other disorders caused by an AIDS weakened immune system is a complication of AIDS. The opportunistic infections associated with AIDS are not restricted to the immune system. Some of the more common are caused by: Pneumocystis pneumonia, Cytomegalovirus, Candidiasis, Mycobacterium tuberculosis and Herpes simplex. AIDS causes also common diseases affecting the gastrointestinal tract, neurological and psychiatric effects, increasing incidence of tumors and cancers (Kaposi's sarcoma, Burkitt's lymphoma, and cervical cancer in HIV infected women due to the human papillomavirus).

About 33 million people worldwide are infected with the virus, with a rate of 2.5 million new infections every year [1]. Over past 30 years, this disease has become a leading cause of mortality worldwide and the main cause of death in sub-Saharan Africa.

There are two types of HIV, each evolved from a different Simian Immunodeficiency Virus (SIV): HIV-1 and HIV-2.

HIV-2 infections are mainly restricted to West-Central Africa although they have been characterized also in India and South Korea. The present work is addressed to the study of the most common type 1 and therefore all the information hereafter reported are related only to this particular family.

HIV-1 is much more prevalent and it is responsible for the AIDS pandemic. Its progenitor is the Simian Immunodeficiency Virus (SIV), which infected chimpanzee

communities in southern Cameroon and was then transmitted to humans. Probably three transmissions took place independently, originating the three large groups into which HIV-1 is nowadays classified: M (*Major*), O (*Outlier*) and N (*Non Major-Non Outlier*). Phylogenetic analysis demonstrated the relation between M and N groups with strains found in chimpanzees, while for group O there are suggestions that probably it originated from gorillas, in which the closest relatives of this group have been identified. In particular the most convincing hypothesis proposes that chimpanzees were the original reservoir of SIVs. Distinct chimpanzee communities in southern Cameroon transmitted divergent SIVcpz to humans, giving rise to HIV-1 group M and N and transmitted HIV-1 group O-like viruses either to gorillas and humans independently, or to gorillas that then transmitted the virus to humans [2].

Group M is the predominant, including the most of the HIV infections worldwide and it has been divided into subtypes denoted with letters (A, B, C, D, F, G, H, J and K) and subsubtypes denoted with numbers (A1, A2, A3, A4, F1 and F2). There are also evidences that intersubtype recombinants are taking part in the pandemic. They result from recombination of subtypes within a dually infected person, and the recombinant form is then passed to other people. These variations are defined as Circulating Recombinant Forms (CRFs), if identified in three or more people with no direct epidemiologic linkage, or Unique Recombinant Forms (URFs), if recovered from only a single person. In Europe, North-America and Oceania, subtype B of group M is the most prevalent one [3].

The epidemiology of HIV-1 is really complex and the classification is ever-changing, trying to follow the rapid evolution of the virus. In fact it has a big genomic diversity, due for example to its high replication rate and to the recombination that, as described before, generates diversity by exchanging long fragments between different genomes. Noteworthy is also the lack of proofreading activity by the reverse transcriptase (RT) that leads to a mutation rate of approximately 3.4×10^{-5} mutations per base pair per replication cycle. The length of the viral genome is about 104 base pairs and every day there is the production of 1010 new virions. Consequently millions of viral variants are produced in a single day within any infected person [3].

HIV-1 is a lentivirus belonging to the retrovirus family. It has a cone shaped capsid with a diameter of 40-60 nm at the wide end and about 20 nm at the narrow end. The capsid is constructed from a lattice of capsid protein (CA) and its most important function is the protection of the genome. Capsid structure must guarantee the stability to survive in the extracellular environment, but also the ability to alter the conformation permitting the release of the genome into the host cell at the appropriate time. A layer of matrix protein (MA) lies between the capsid and the envelope. There are two proteins, essential for the recognition and internalization steps, associated to the envelope at the surface of the virion. A transmembrane protein (TM) called gp41 (glycoprotein with a molecular weight of 41 kD), that presents the C-terminal bound to the matrix protein inside the virion, and a heavily glycosilated surface protein (SU) called gp120 (glycoprotein with a molecular weight of 120 kD). A trimer of gp41-gp120 is called spike and it is possible to see about 14 spikes at the surface of each virion (Figure 1.1).

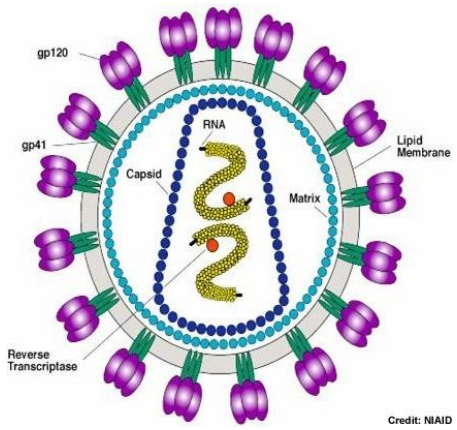


Figure 1.1 HIV-1 virion structure.

HIV-1 contains two copies of single stranded RNA present as a dimer formed by base pairing between complementary sequences. The length of the genome is about 9.3 kb (Figure 1.2) and it encodes the *gag*, *pol* and *env* genes, together with auxiliary genes known as *Tat*, *Rev*, *Nef*, *Vif*, *Vpr* and *Vpu*, important for gene expression, transport of components within the cell and modification of the host's immune response. At the ends of the described genes, there are long terminal repeated sequences are known as LTR. LTR include promoters, enhancer, origins of replication and other elements involved in the control of the transcriptional regulation like the Transactivation Response Element TAR, which directly interacts with the Tat protein [4, 5].

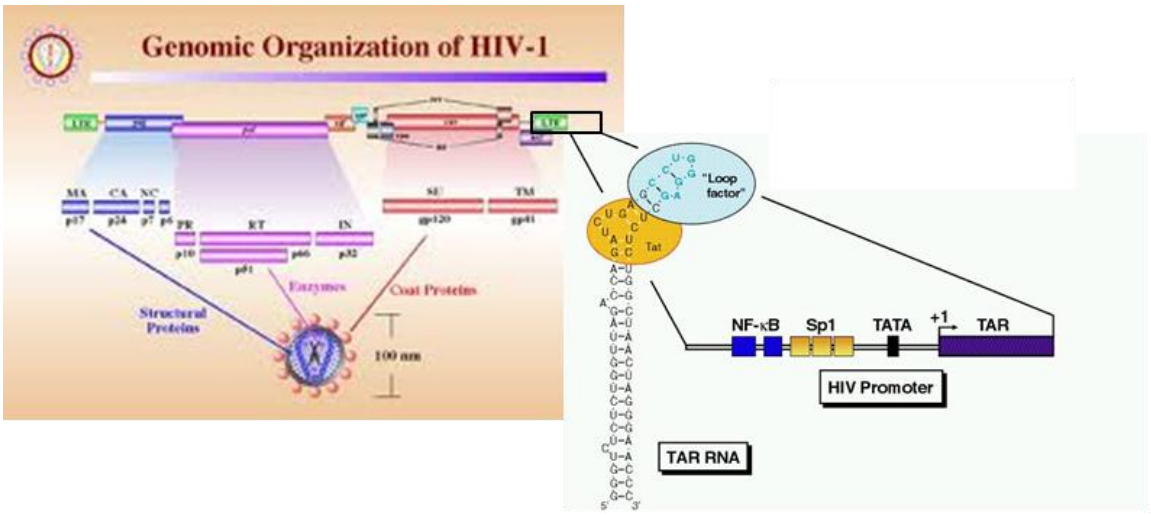


Figure 1.2 Genomic organization of HIV-1. Schematic representation of the genes encoding for structural protein, viral enzymes and coat proteins; focus on LTR, long terminal repeated regions.

It is possible to find proteins bound non-covalently to the nucleic acids. Usually these proteins are encoded by the virus genome. The most abundant is the nucleocapsid protein (NC). NC is rich in basic amino acids (29% of the residues are lysine, arginine or histidine) and have two zinc fingers. Noteworthy is also the presence of a molecule of transfer RNA ($tRNA^{lys-3}$) of the host cell, bound of each copy of the viral RNA through base pairing in a specific area called Primer Binding Site (PBS). This tRNA, together with other molecules of the host cell, is packaged in the virion during assembly. It is the primer for the synthesis of the (-)DNA by reverse transcriptase.

1.1.2 HIV-1 Replication

HIV can only replicate inside human cells and begins its cycle by entries those cells that present CD4 receptors on their surface. This receptor is usually found on immune cells such as T cells and macrophages. The gp120 surface protein interacts with the CD4 marker resulting in attachment of the virus to the host cell. However, for entry, the virus also needs a co-receptor in addition to the CD4 receptor i.e. CCR5 (R5) on macrophages and CXCR4 (X4) on T cells. The interaction between the CD4-gp120 complex and the co-receptor is required for changes in the cell membrane to facilitate viral entry.

Figure 1.3 summarizes all the HIV life cycle steps.

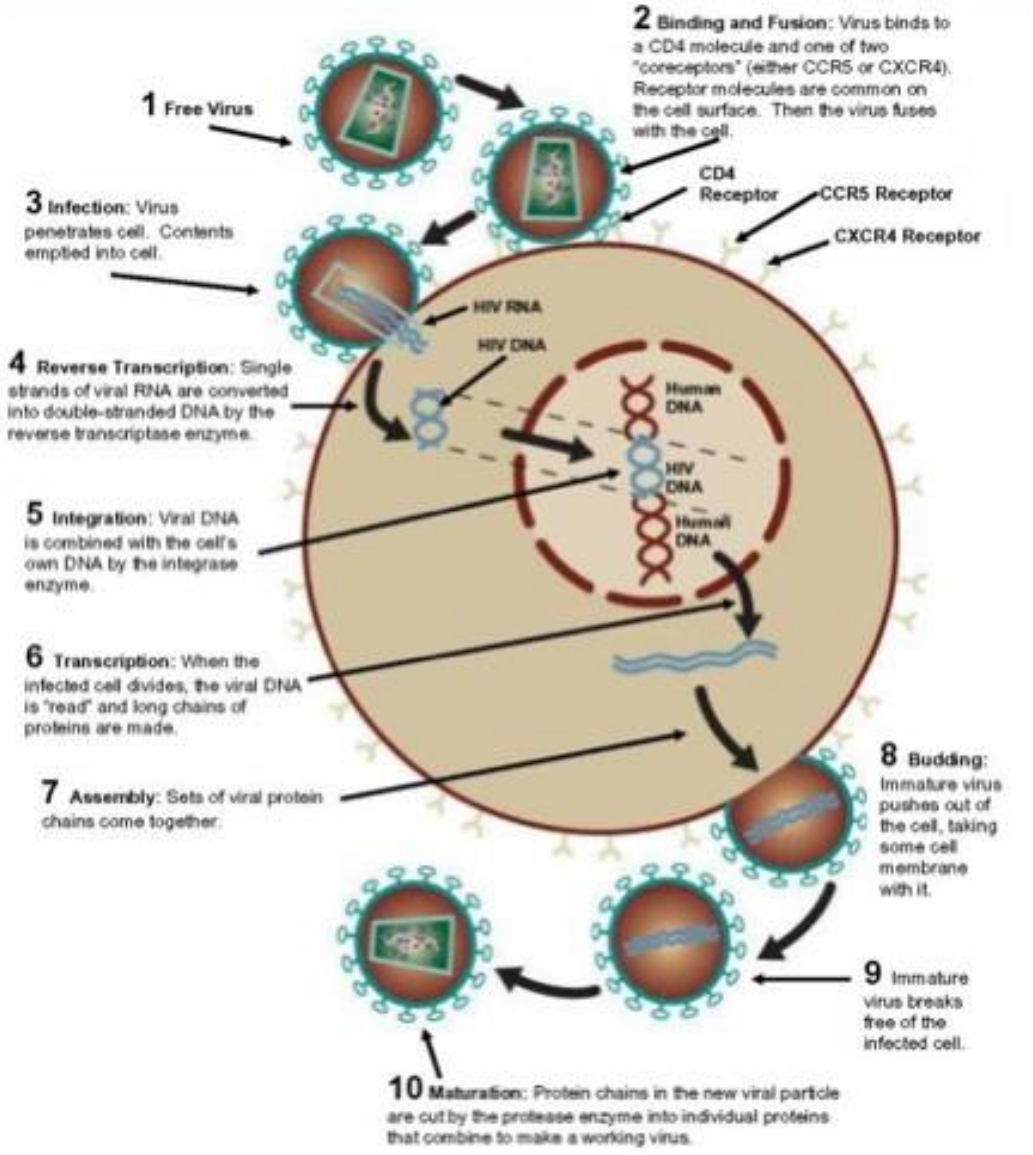


Figure 1.3 Schematic representation of HIV-1 life cycle steps.

In particular we focus on Reverse Transcription, in which NC protein is involved, and on Transcription that is a post-integration step Tat-mediated.

Reverse Transcription: The single stranded viral RNA must be converted to double stranded DNA by the viral enzyme reverse transcriptase, to enable HIV to be integrated into the host DNA and to use the cellular genetic machinery to make new viruses. The primer for synthesis of the minus strand DNA is the tRNA bound to the PBS (Primer Binding Site) of the viral genome, while a 19 nucleotide long polypurine tract (PPT) near the viral 3' end is the primer site for the synthesis of the plus strand DNA. The DNA that results from transcription is longer than the RNA genome, because contains LTRs (long terminal repeats) at each ends. Detailed information about reverse transcription and NC implications in reverse transcription are reported on paragraph 1.1.6.

Transcription. Once the viral DNA is integrated into the host genome, the virus uses the transcription machinery of the infected cell to generate RNAs that will function as mRNAs or as genomes for progeny virions. The two integrated LTRs of the provirus are functionally different, even presenting identical sequences. Cell transcription factors bind to promoter and enhancer sequences in the upstream LTR, and transcription initiates at the U3-R junction by cell RNA polymerase II. Transcription is then terminated in the downstream LTR, at the RU5 junction after polyadenylation signal. Many of the genome-length transcripts are spliced and mainly three size classes of virus transcript can be detected in infected cells: the largest is of entire genome length (9.3 kb), one derives from a single splicing event (around 4.5 kb) and the final class, of about 2 kb, is the result of multiple splicing. All together, the splicing events result in more than 30 mRNA species. Each transcript is capped and polyadenylated and afterwards exported from the nucleus [6].

Tat protein (Transactivator of Transcription) enhances the transcription after binding to the TAR region (Transactivation Response Element) at the 5' end of the nascent transcripts. In particular it has a nuclear localization signal that directs it to the nucleus and, as other cell kinases (for example CycT and Cdk9), binds to the structured TAR element. The consequent phosphorylation of the cytoplasmic tail of RNA Polimerase II, results in the release of stalled enzyme and allows its rapid elongation along the proviral DNA, with increased rate of transcription. In absence of Tat, transcripts are incomplete but they are enough to allow the synthesis of a small amount of Tat, which then assists the synthesis of genome-length RNA.

1.1.3 Pharmacological AIDS Treatments and Drug Resistance

Several drugs with different targets and mechanisms of action had been approved since HIV-1 was first discovered. The classes of drug which have been approved since now are reported below and in Figure 1.4.

FDA-approved therapies for the treatment of HIV – October 2007.

Approval date	Generic name	Brand name	Manufacturer
Nucleoside reverse transcriptase inhibitors (NRTIs)			
1987	Zidovudine (AZT)	Retrovir	GlaxoSmithKline
1991	Didanosine (ddI)	Videx	Bristol-Myers Squibb
1992	Zalcitabine (ddC)	Hivid	Roche Pharmaceuticals
1994	Stavudine (d4T)	Zerit	Bristol-Myers Squibb
1995	Lamivudine (3TC)	Epivir	GlaxoSmithKline
1997	Lamivudine + zidovudine	Combivir	GlaxoSmithKline
1998	Abacavir	Ziagen	GlaxoSmithKline
2000	Abacavir + lamivudine + zidovudine	Trizivir	GlaxoSmithKline
2000	Didanosine (ddI)	Videx EC	Bristol-Myers Squibb
2001	Tenofovir disoproxil fumarate	Viread	Gilead Sciences
2003	Emtricitabine (FTC)	Emtriva	Gilead Sciences
2004	Abacavir + lamivudine	Epzicom	GlaxoSmithKline
2004	Emtricitabine + tenofovir	Truvada	Gilead Sciences
Non-nucleoside reverse transcriptase inhibitors (NNRTIs)			
1996	Nevirapine	Viramune	Boehringer Ingelheim
1997	Delavirdine (DLV)	Rescriptor	Pfizer
1998	Efavirenz	Sustiva	Bristol-Myers Squibb
Protease inhibitors			
1995	Saquinavir	Invirase	Roche Pharmaceuticals
1996	Ritonavir	Norvir	Abbott Laboratories
1996	Indinavir (IDV)	Crixivan	Merck
1997	Nelfinavir	Viracept	Pfizer
1997	Saquinavir mesylate	Fortovase	Roche Pharmaceuticals
1999	Amprenavir	Agenerase	GlaxoSmithKline
2000	Lopinavir + ritonavir	Kaletra	Abbott Laboratories
2003	Atazanavir	Reyataz	Bristol-Myers Squibb
2003	Fosamprenavir	Lexiva	GlaxoSmithKline
2005	Tipranavir	Aptivus	Boehringer Ingelheim
2006	Darunavir	Prezista	Tibotec, Inc.
Fusion inhibitors			
2003	Enfuvirtide (T-20)	Fuzeon	Roche Pharmaceuticals & Trimeris
Multiclass combinations			
2006	Efavirenz + emtricitabine + tenofovir disoproxil furmarate	Atripla	Bristol-Myers Squibb and Gilead Sciences
Entry inhibitors			
2007	Maraviroc	Selzentry	Pfizer
Integrase inhibitors			
2007	Raltegravir	Isentress	Merck

Figure 1.4 FDA approved anti HIV-1 drugs divided depending on their molecular target. From [7].

The most successful anti HIV-1 treatment is known as combined antiretroviral therapy or highly active antiretroviral therapy (HAART). HAART has been the major accomplishment for the chronic suppression of HIV-1 replication in patients since 1996 and allowed AIDS to become a treatable disease.

HAART works remarkably well at controlling the virus, but unfortunately, HIV-1 cannot be eradicated completely due to the existence of the latent viral reservoirs [8, 9]. Usually, three different drugs from at least two of the previous groups are taken together as this greatly delays the development of drug resistant viral variants, by targeting multiple steps of the HIV-1 life cycle. Nearly all drugs currently used to treat HIV eventually stop working. This has resulted from HIV evolving in response to selection from antiretroviral drug treatments. Mutations are common in HIV because of the rapid rate at which it replicates and because it lacks of proofreading mechanisms. This is why additional therapeutic approaches are necessary and why research is working with the aim to discover novel inhibitors of HIV-1 targeting other stages of the viral life cycle.

During the last few years our attention was focusing on the emerging target NC protein. The inactivation of NC protein, in order to reduce the reverse transcription of the viral genome, could provide a new anti-HIV therapy.

Moreover, a promising class of anti-HIV agents is represented by compounds which interfere with the Tat-mediated transcription process of the viral replicative cycle.

1.1.4 HIV-1 Reverse Transcription

Reverse transcription is a critical step for retroviral replication. Reverse transcription is catalyzed by Reverse Transcriptase (RT) and consists of a complex series of events that culminate in the synthesis of a linear double-stranded DNA copy of the viral RNA genome. The conversion of a single-stranded RNA genome to a double-stranded DNA is achieved by employing two obligatory strand transfer steps, minus-strand transfer and plus-strand transfer. During the assembly of viruses, both host $tRNA^{Lys-3}$ and viral RNA are selectively packaged into the virions.

In figure 1.5 is reported the mechanism of reverse transcription [10].

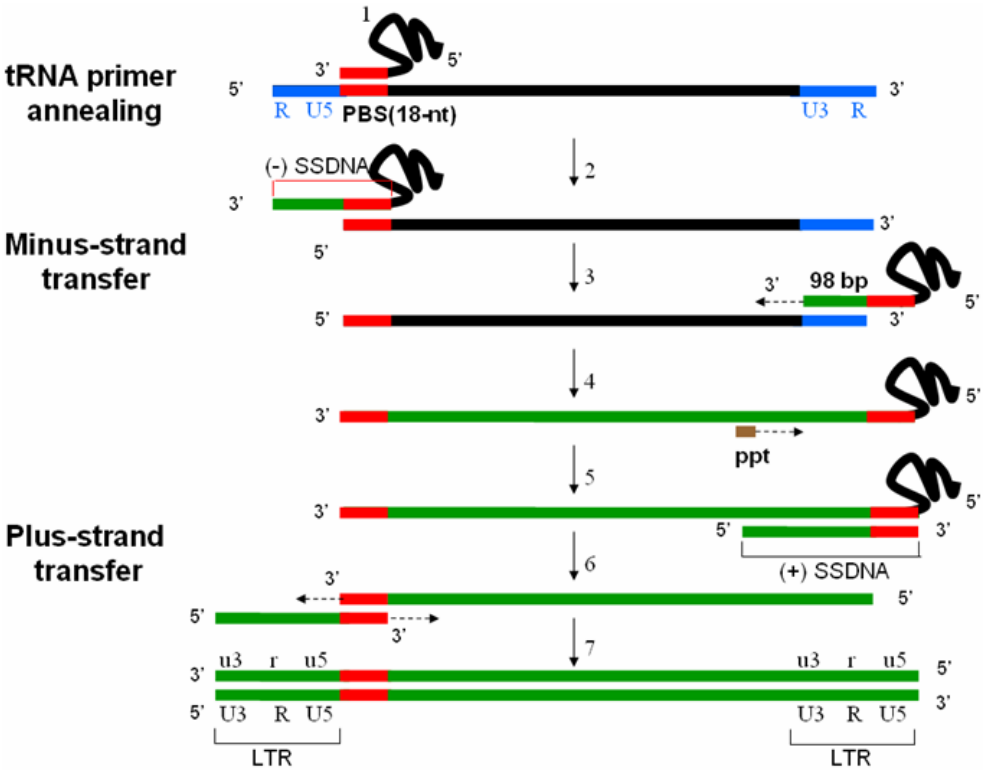


Figure 1.5 Mechanism of reverse transcription. Step 1: Annealing of $tRNA^{Lys-3}$ to the PBS. Step 2: Synthesis of (-) SSDNA and digest of RNA template by RT. Step 3: Minus-strand transfer. Step 4: Elongation of minus-strand DNA to the 3' end. Step 5: Synthesis of (+) SSDNA initiated from PPT as a primer. Step 6: Plus-strand transfer. Step 7: Elongation of both DNA strands. Abbreviations: U5/3, 5'/3' untranslated region; PBS, primer binding site; SSDNA, strong stop DNA; PPT, polypurine tract; LTR, long terminal repeat region. [10]

1. The HIV-1 reverse transcriptase initiates proviral DNA synthesis from a $tRNA^{Lys-3}$ annealed to an 18-base sequence of the primer binding site (PBS) near the 5' end of the RNA.
2. Upon the hybridization of $tRNA^{Lys-3}$ to PBS, the polymerase domain of viral RT copies the RNA into DNA sequences, and simultaneously the RNase H domain of RT degrades

the genomic RNA which has been copied. Synthesis of the DNA copy proceeding to the 5' end of the genomic RNA generates the minus-strand strong stop DNA ((-) SSDNA).

3. In order to complete the reverse transcription, (-) SSDNA must be transferred to the 3' end of the RNA genome. Since the genomic RNA has two identical R regions in the terminus, base-pairing (-) SSDNA (containing partially sequences of the R region) with the RNA genome can be switched from the 5' end to the 3' end.

   This so-called minus-strand transfer step is obligatory and is either intra-molecular (i.e. from one end to the other end of a genomic RNA) or inter-molecular (i.e. from one to the other copy of genomic RNA).
4. After the minus-strand transfer, RT continues to elongate the same DNA transcript until it reaches the polypurine tract (PPT) site located near the 3' end of the viral genome.
5. The RNase H domain of RT makes a specific nick at the PPT of 5' end of RNA genome and creates an RNA primer to initiate plus-strand DNA synthesis.
6. Once the synthesis of plus-strand DNA reaches the 5' end of minus-strand DNA, plus-strand strong-stop DNA is then transferred to the other end of the almost complete minus-strand DNA, presumably by the complementarity between the PBS.
7. This obligatory plus-strand transfer is primarily intra-molecular and enables the elongation of plus-strand DNA to proceed to the 3' terminus. Two obligatory transfers result in extension of both DNA transcripts and create the long terminal repeat regions (LTR) containing extra untranslated regions that do not exist in the viral RNA at both termini.

1.1.5 TAR (Trans-Activation Responsive Element)

TAR is an RNA sequence that is highly conserved in HIV-1 and is characterized by a particular secondary structure.

Various elements may be present within an RNA secondary structure (Figure 1.6):

- *stem,* that is a double stranded RNA region in which the pairs of paired nucleotides are stacked one upon the other to form the double helix (in the form A);
- *internal loop,* that is a single stranded region in both strands within the RNA helix;
- *bulge,* that is a region, within the RNA helix, in which in only one strand there is one or more bases unpaired;
- *hairpin loop,* that is a single stranded region at the terminal portion of the RNA helix;
- 5' *overhang* that is a 5' unpaired sticky end.

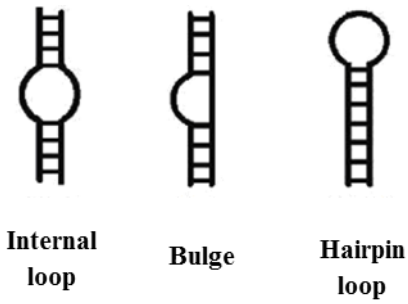


Figure 1.6 Schematic representation of element charachterizing RNA secondary structures.

In particular, TAR is a 59-mer *stem-loop* structure characterized by the presence of a *loop* (6 nucleotides) separated from the *bulge* by a double stranded helix (*stem*). The *bulge* is constituted by 3 nucleotides UUU or UCU (Figure 1.7).

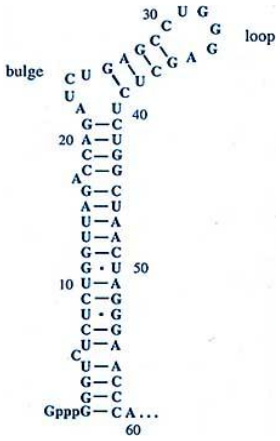


Figure 1.7 Sequence and structure of TAR RNA.

TAR plays a very important role in the HIV-1 life cycle: it is involved in fundamental steps like reverse transcription and transactivation.

1.1.6 The HIV-1 nucleocapsid (NC) protein

NC structure

Shortly subsequent to virion budding, the virion undergoes a major structural change termed maturation in which the Gag polyprotein is cleaved by the virus-encoded protease (PR) into a series of products, including nucleocapsid (NC) proteins.

The HIV-1 NC is a zinc-finger protein with 55 amino acids. NC is composed of two zinc-binding domains of the CCHC type flanked by a flexible polypeptide chain and connected by a short basic peptide linker (Figure 1.8) [11-14]. About 2000 copies of NC are tightly

associated with the genomic RNA in the core of mature virions [15, 16] and are believed to protect the genomic RNA against degradation by RNase [17].

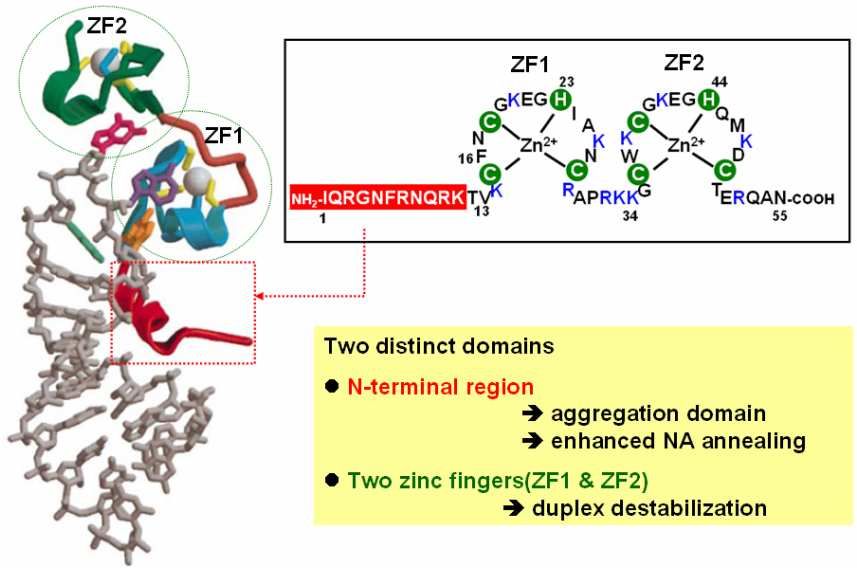


Figure 1.8 Three-dimensional NMR structure of a NC/RNA hairpin complex. The N-terminal 12 residues of NC (shown in red) bind to the major groove of the RNA stem (shown in gray) through the electrostatic interactions. The two zinc fingers (ZF1 and ZF2, shown in blue and green) engage in specific interactions three purine bases (shown in green, pink, purple and orange) in the tetraloop. The zinc ions are shown in white. **(Inset)** The primary structure of NC protein shown is from the NL4-3 isolate of HIV-1. The cysteine and histidine residues coordinating with zinc highlighted in green show both zinc fingers belonged to the CCHC type. The basic residues are shown in blue. This figure is adapted from B. G. Turner and M. F. Summers, 1999 [16].

The strict conservation of the CCHC zinc-binding motif in all retroviruses containing NC ZFs and the absence of other common zinc binding motifs, such as the CCCC or CCHH sequences, strongly suggested a critical role for these structures in multiple events during virus replication. Moreover, the absolute conservation of the CCHC motif together with the fact that all the known mutations in this motif are lethal for the virus, raises the possibility that HIV-1 will be unable to generate mutants that are resistant to drugs [18].

NC binds nucleic acids

The function of NC relies on two interactions, electrostatic and π- π stacking interactions, with nucleic acids. NC protein appears to have two distinct domains implied in binding with nucleic acids.

The N-terminal basic helix of NC interacts nonspecifically with the major groove of the RNA stem by electrostatic interactions [10] and is responsible for the aggregation of NC/nucleic acid complexes [19, 20].

The zinc finger domain interacts with specific base sequences in particular with a preference for TG- or UG- rich regions [21]. The binding of zinc fingers to nucleic acids NA is responsible for destabilizing their secondary structure. NMR studies further showed that Phenylalanine in the second zinc finger (ZF1) and Tryptophan in the second zinc finger (ZF2) engaged in aromatic π- π stacking interaction with the G residues in the single-stranded GGAG tetraloop as shown in Figure 1.8. NC was observed to have highly specific interactions with bases in the nucleic acids loops since binds specifically to exposed (unpaired) nucleobases of RNA or DNA.

NC chaperone activity

NC is crucial for many processes of the viral life cycle in which interactions between nucleic acid and NC are the key. For example, NC is involved in the recognition and packaging of RNA genomes, in virus assembly, in selectively packaging $tRNA^{Lys-3}$ primer and in many annealing steps during the reverse transcription [10].

As a nucleic acid chaperone protein, NC catalyzes nucleic acids conformational rearrangement that leads to the most thermodynamically stable structure. The chaperone activity of NC is believed to be derived from two main consequences of the nucleic-acid/NC interactions in this system.

First, NC lowers the energy of highly structured nucleic acids by the complexation with nucleic acids and further partially melting nucleic acids.

Second, NC lowers the energy cost of bringing two complementary strands of nucleic acids together by screening the negative charges of two oligonucleotides [20-22] and perhaps through the binding of the N-terminal basic helix to the double-stranded stem of the hairpins.

The NC chaperone activity allows the protein to be a cofactor of RT and promote several annealing reactions during the HIV-1 reverse transcription process: NC makes it possible for RT to copy the stem-loop sequences in RNA and DNA templates by melting the structured sequences. In additions, NC has been shown to facilitate the annealing of cellular tRNA onto the PBS of the viral RNA, the annealing of complementary PBS sequences in plus-strand transfer and the annealing of the transactivation response (TAR) region of the viral RNA to the complementary sequences (cTAR) during minus-strand transfer [10].

1.1.7 NC Role in the Minus-Strand Transfer

The minus strand transfer consists in the transfer and in the annealing of the (–) SSDNA to a complementary R region at the 3′ end of the RNA template. The result is for DNA/RNA hybrid to lose 18 base-pairs and to gain 98 base-pairs. In detail, the annealing occurs between the trans-activation responsive region (TAR) of the viral RNA and the complementary sequences (cTAR) at the 3’ end of the (–) SSDNA, both contained in the repeat regions R (Figure 1.9).

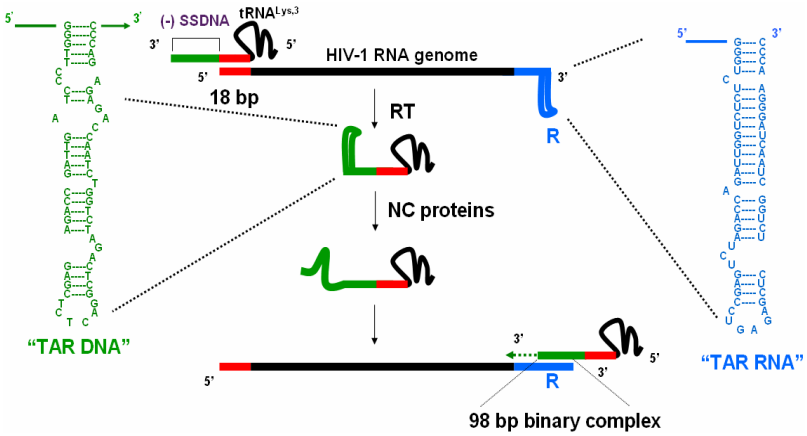


Figure 1.9 NC key role in the minus strand transfer. NC is shown to destabilize the highly structured TAR RNA (shown in blue) and cTAR DNA (shown in green) and to promote the annealing of these two viral regions leading to the formation of a 98-nt base pair heteroduplex. From Hsiao-Wei Liu, Single-Molecule Studies on the Role of HIV-1 Nucleocapsid Protein/Nucleic Acid Interaction in the Viral Replication Cycle, 2007 [10].

However, although thermodynamically favored, this reaction does not occur significantly in the absence of NC that has different activities:

- NC binds to TAR (RNA) and cTAR (DNA) and melts their secondary structure,
- NC inhibits the self-priming and
- NC promotes the TAR/cTAR hybrid formation.

- cTAR and TAR regions are highly structured regions containing double-stranded segments separated by numerous conserved bulges, mismatches and an internal loop. Formation of these secondary structures results to be incompatible with successful strand transfer and with an efficient synthesis of DNA. NC protein destabilizes these stem-loop structures. Melting, as schematically represented in Figure 1.10 is thought to occur preferentially at locations where the stem may be less stable, as for example at the level of the two short, double-stranded segments at the stem terminus.

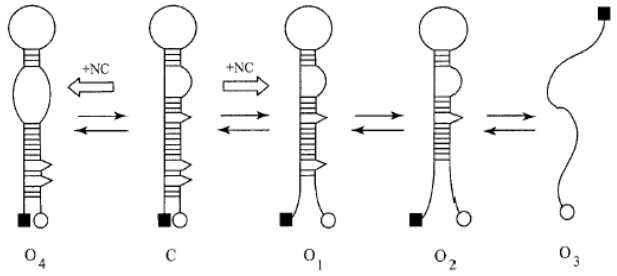


Figure 1.10 Proposed scheme for the melting of cTAR secondary structure. In the absence of NC, the closed species (C) is favored, though it is already in equilibrium with partially or totally (O1 to O4) melted species. NC shifts the equilibrium toward the melted species. The square and the circle correspond to the 5' and 3' ends respectively [23].

- The cTAR stable intramolecular secondary structures enable the DNA to fold back on itself and self-prime. NC was seen to inhibit the competing intramolecular self-priming of (–) SSDNA [24] that occurs due to the presence of the cTAR initiating unwanted elongation by the RT (Figure 1.11).

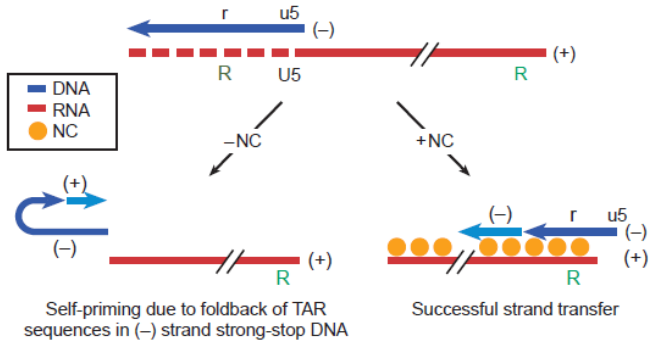


Figure 1.11 In the absence of NC protein a competing reaction occurs: self complementary sequences can cause the - SSDNA to fold back on itself and prime. Addition of NC to reverse transcription prevents this self priming reaction and increased the level of the synthesis of full-length strand DNA. From [25].

- The modification of the cTAR and TAR structures that governs the annealing reaction is found strictly dependent on the integrity of the ZFs hydrophobic platform [26]. The two Zn^{2+} fingers are required to ensure specificity and fidelity of viral DNA synthesis. cTAR/TAR annealing reaction in the absence of NC is slow, depending on nucleation through the apical loops and proceeds via a well-defined two-steps pathway. The first step involves fast, reversible formation of an unstable intermediate that is likely to be an extended kissing complex wherein loop, upper stem and bulge nucleotide are base-paired. This intermediate converts into an extended duplex [27] through a slow strand exchange between the terminal stems. NC switches the reaction pathway from the poorly efficient loop-loop pathway to a highly efficient zipper pathway (Figure 1.12) by directing the hybridization of these

sequences through the 3'/5' termini of their double stranded stems and by forming a 'zipper' intermediate.

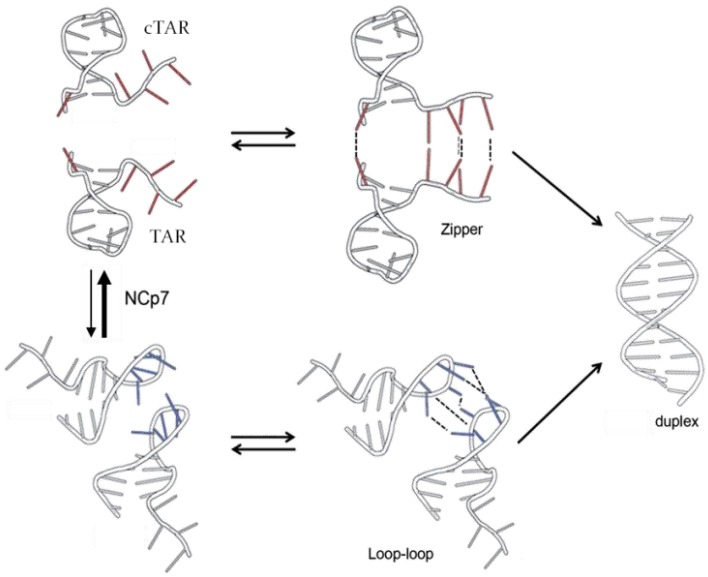


Figure 1.12 Stem-initiated annealing involves nucleation through the 3′/5′ termini resulting in formation of a "zipper" intermediate followed by conversion to the fully annealed product and occurs in the presence of NC. Adapted from [26].

The TAR/cTAR annealing reaction catalyzed by NC is essential for the progression of reverse transcription. Compounds able to inhibit the hybrid formation could damage the HIV life cycle and could be considered potential anti-HIV agents.

1.1.8 Anti-NC Strategies

The inactivation of NC protein, in order to reduce the reverse transcription of the viral genome, could provide a new anti-HIV therapy. Drugs targeting NC protein are currently not available, however different classes of molecules have been proposed as anti-NC agents as reported below.

- *Peptidomimetics*: an alternative strategy to inhibit NC functions is to design peptides that can directly compete with NC for the binding to its RNA and DNA substrates i.e. cyclic peptides mimicking the spatial orientation of Phe16 and Trp37 residues in the hydrophobic plateau and containing two basic residues to mimic Arg26 and Arg32. However, the activity of the leading compound is not so satisfactory. Attempts are actually done to better mimic the NC hydrophobic plateau, in order to increase the activity of these peptidomimetics *i.e.* with small peptides able to specifically recognize NC targets that are composed of a cluster of Trp residues surrounded by basic residues. These peptides were able to compete with NC target

sequences binding mainly through stacking interactions between the Trp residues and the oligonucleotide bases and to stabilize the cTAR secondary structure inhibiting the NC-directed melting of the cTAR sequence [28].

- *RNA aptamers* were selected using the SELEX method, obtaining aptamers of about 40 nucleotides in length. Their secondary structures correspond to stem-loops in which the stems are rich in GC base pairs and the loops contain G and U residues [29].
- *Zinc ejector agents* are oxidizing agents that are able to attack the sulfhydryl group of the Cys residue and eject the zinc, rendering the zinc fingers and by extension, the whole protein, inactive. Many classes of Zinc ejector agents and zinc ligands have been identified: NOBA (3-nitrosobenzamide), DIBA (disulfide benzamide) [30]. However, these compounds elicit serious side effects since they target cellular proteins binding zinc and this impairs their use in therapy.
- *Non zinc ejecting NC binders*: antagonists of the strong binding of NC to oligonucleotides were identified. Many of them contain a xanthenyl ring structure while other had a related fluorescein or gallein-like structure (containing a xanthene ring substituted by two hydroxyls in positions 4' and 5', lead compound tetrachlorogallein). Hydroxyls groups likely play a critical role since they were found to be critical for binding to NC and are the only feature that the compounds share. Probably hydroxyl groups bind to the amide nitrogen of Gly35 with additional contacts at the carbonyl oxygens of Gly40 and Lys33 of NC and it is thought that these compounds might interfere with NC's ability to stack its aromatic residues with the bases of its nucleic acids target [31]. Compounds able to inhibit NC activity by binding do not show any structure-activity correlation.
- *Intercalators*: Actinomycin D (widely used anti-cancer drug that inhibits the DNA-dependent DNA/RNA synthesis through DNA intercalation and binding to the minor groove) was the most active, but the main problem is that it interacts with a large range of nucleic acids, giving a poor specificity. As a consequence, a clinical use of such a molecule will likely generate critical side effects and the emergence of resistant strains [32].

1.1.9 Intercalating TAR-binders

A wide range of small molecules with different chemical structures from our in-house chemical library were analyzed in our lab: in this work we focus on molecules with intercalating activity or able to interact by stacking with the base pairs of nucleic acids. Although the regions in which the RNA double helix is interrupted by single chain sequences

proved to be optimal for interactions with small molecules, also the stem of RNA can be considered as target for compounds.

Several studies reported in literature describe the ability of classic DNA intercalating agents, in particular, ethidium bromide and proflavine, to bind with the same mechanism also the TAR-RNA [33]. These results encouraged the analysis of other agents capable of binding the double strand region of TAR: Hoechst 33285 (Figure 1.13) was synthetized with the aim to make the binding selective for the bulge of the TAR.

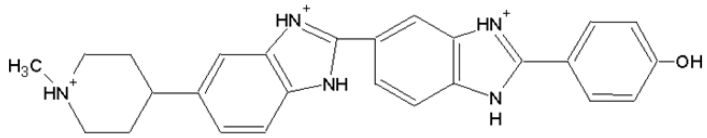


Figure 1.13 Chemical structure of Hoechst 33285.

Other compounds characterized by a heterocyclic region were developed for their ability to give stacking with the bases of the bulge of TAR, linked to a positively charged portion capable of binding to the RNA phosphate backbone. By varying the length and composition of the linker or the strength of the electrostatic component, it was possible to modulate the efficacy of the compounds [33].

Moreover, ethidium-arginine conjugats were synthetized: they proved to be able to associate the ability of ethidium to form stable complexes with RNA, with the ability of arginine to interact at the level of the bulge. This was highlighted by the study of argininamide, a small molecule capable of binding to TAR through electrostatic interactions between its guanidinium group and the phosphate groups of residues U23 and A22 of the TAR bulge.

This binding stabilizes the conformation of free TAR, changing the average angle between the two double-stranded regions adjacent to the bulge. The conjugates ethidium-arginine are able to bind strongly and specifically the TAR in the inactive form, thus interfering with the formation of the Tat-TAR complex: arginine-RNA interaction occurs in correspondence of nucleotides 23, 24, 25 of the bulge, while the intercalation takes place between the pair of bases C18-G44 (Figure 1.14).

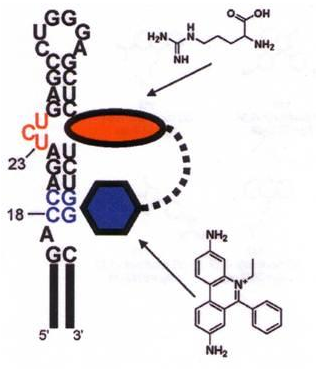


Figure 1.14 Schematic representation of the ethidium-arginine conjugates binding to TAR (PIC).

Finally, another class of developed compounds is represented by *Threading Intercalators* that are intercalating agents characterized by a polyaromatic nucleus disubstituted at opposite positions with charged side chain. The intercalating mechanism of these compounds is based on the heterocyclic portion between the bases of the double-stranded RNA by positioning the two side chains in the two different grooves of the helix. This process requires the passage of the molecule through the duplex and therefore the opening of the double helix at the binding site. It therefore takes place preferentially in dynamic regions of RNA and it is for this reason that the TAR, with its flexible bulge-loop structure, is a good target for these molecules.

The peptide-intercalators (PIC: Peptide-Intercalator Conjugate) are small molecules able of binding to the double stranded RNA with high affinity and selectivity, according to the intercalating mechanism shown in Figure 1.15.

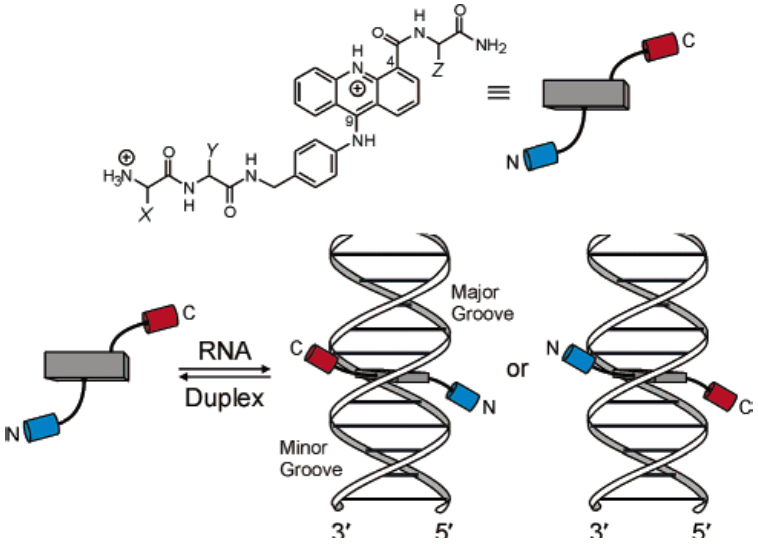


Figure 1.15 Schematic representation of the *Threading Intercalator* process of *Peptide-Intercalator Conjugate*.

Intercalating agents that can selectively bind to nucleic acids (TAR, cTAR and the TAR/cTAR hybrid) involved in the annealing steps catalyzed by NC (TAR, cTAR and the TAR/cTAR hybrid) could be developed as potential anti-NC compounds able to interfere with the reverse transcription process.

1.1.10 The transcriptional activation process

Transcriptional activation of the HIV-1 long terminal repeat (LTR) promoter element by the viral Tat protein is an essential step in the HIV-1 life cycle. This process represents a promising anti-HIV target, because not only it is directly involved in the replication process of HIV-1, but also because it implies highly conserved viral elements as the Trans-Activation Response element TAR and the peptide Tat.

The interaction of Tat with cellular factors, such as the cyclin-dependent kinase CDK9 and Cyclin T1, is a fundamental event for the recognition with high affinity of TAR, that determines the increase of the RNA polymerase II processivity. Otherwise, the transcriptional elongation of the viral mRNA would end immediately after the synthesis of short transcripts. In the absence of Tat, transcription normally starts, but loses efficiency because the promoter recruits poorly processive polymerases that detach prematurely from the template DNA (Figure 1.16).

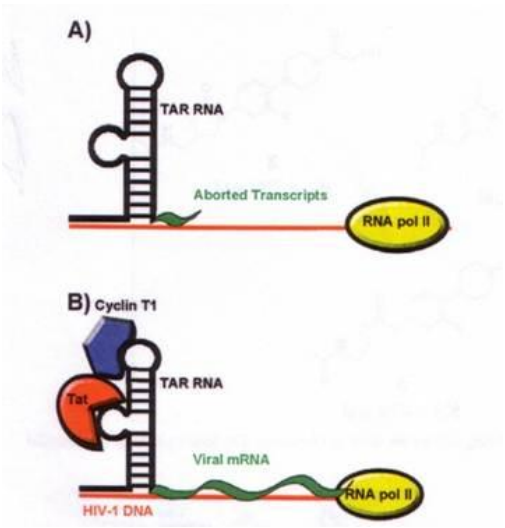


Figure 1.16 The Tat-TAR interaction enhances the RNA polyII processivity. (A) In the absence of Tat, RNA polyII (yellow) synthetizes short fragment of not-polyadenilated RNA (green). (B) The Tat (red) binding to TAR determines the recruitment of other protein including Cyclin T1 (blue). The all complex enhances the efficiency of the RNA polyII.

The transcription process can be summarized as follows: the early steps are regulated by numerous cellular factors, called basal transcription factors (TFII D-B-F-E-H), which, by interaction with specific sequences at the level of the viral promoter, form, together with the RNA polymerase II, the "competent" elongation complex [34]. The phosphorylation of the carboxyl-terminus of the polymerase by the CDK7 kinase, a component of the initiation factor TFIIH, allows the clearance of the promoter and the initiation of TAR transcription. TAR,

assuming its characteristic stem-loop structure, recruits Tat protein in the transcription complex.

Also the cellular cofactor TAK, a Tat-associated kinase, is involved in the Tat-TAR interaction with the formation of the ternary complex Tat-TAR-TAK. This cofactor, identical to P-TEFb that is a positive factor of transcriptional elongation [35], contains the kinase subunit CDK9 and the kinase-dependent Cyclin T1 [36]. CDK9 phosphorylates the carboxyl-terminal domain of the RNA poly II, allowing an efficient processivity.

1.1.11 Tat protein in HIV-1

Tat, a potent transcription activator, is an arginine rich protein constituted by 101 aminoacids, well conserved in lentiviruses. Properties that distinguish Tat from other transcription factors include that it acts predominantly at the level of transcription elongation, rather than initiation, and that the target sequence for Tat is an RNA stem–loop structure, TAR, located immediately 3' to the LTR transcription start site [37].

Tat is encoded by two exons: the amino acids encoded by the first exon (amino acids from 1 to 72) are necessary both for binding to the mRNA target sequence and for the transcription activation; the second exon contains the information necessary to activate the different pathogenic functions of HIV and the immunosuppressive activity of Tat (Figure 1.17).

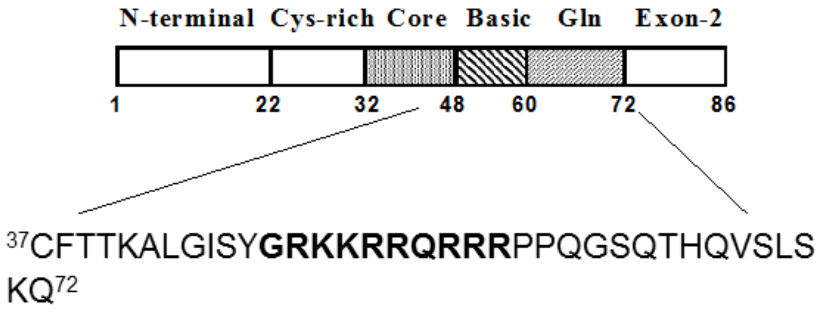


Figure 1.17 Schematic representation of HIV-1 Tat protein.

It was shown that 86 amino acids are sufficient to maintain all functions of this molecule, including the propagation of HIV in vivo [38] and this minimum active portion is divided into 6 distinct domains (Figure 1.17):

- **1-21:** N-terminal domain characterized by hydrophilic and hydrophobic residues arranged in a periodic manner. The high concentration of proline in this region can have the role of preventing the degradation of Tat by proteases;

- **22-37:** domain rich in cysteine residues, which form a scheme well conserved in different subtypes of HIV-1. It has been experimentally observed that the deletion of this region leads to a reduction of the transactivation activity of Tat, probably because this region is involved in interactions with a number of cellular co-factors;
- **38-48:** extremely conserved "core" of 10 hydrophobic amino acids. Noteworthy is the lysine residue 41, whose deletion leads to the loss of TAR recognition with consequent loss of transactivation;
- **49-59:** basic domain, containing 6 arginine *(Arginine Rich Region),* and 2 lysine, characteristic of the RNA-binding proteins and responsible of the interaction with TAR; the deletion of this domain leads to the loss of Tat activity. This region is moreover involved in Tat ability to go through cellular and nuclear membranes;
- **60-*72*:** glutamine rich, less conserved region, which forms a α-helix when Tat binds to TAR.

1.1.12 Tat-TAR interaction

A lot of studies contributed to understand the characteristics of the Tat-TAR interaction: the protein binds directly to U23, U38 and U40 bases of the RNA major groove and, during the interaction, the protein is located in close proximity to the O6 of G21 and G26 in the RNA major groove [39, 40]. The Tat basic region is directly involved in the interaction with TAR. Using deletion assays, the 38-72 region, containing the basic region of the protein, resulted to be able to bind to TAR with affinity comparable to that of the full-length Tat [41], while the 48-57 region represents the necessary and sufficient portion to bind to the bulge of TAR. In particular, lysines contribute to the complex formation through electrostatic interaction; arginines, arginine 52 in particular, determine the functional TAR conformation [42]. Finally, the cysteine rich domain seems to be implicated in the binding with the cellular protein interacting with the loop of the nucleic acid. Cross-linking assays allowed determining that the amino acid 57 of Tat, during the complexation with TAR, is located in close proximity to the nucleotide 31 of the TAR loop [40].

Without Tat, the tree bases forming the TAR bulge show a high flexibility that allows a number of different conformations of the bulge, but only the Tat-induced conformation results to be functional.

Molecular modeling studies show the conformational variation of TAR in the presence and in the absence of Tat (Figure 1.18): without Tat, the U23, C24 and U25 bases are stacked on A22 and form the bulge; in the presence of Tat, the bulge results to be altered i.e. U23

moves to G26, and C24 and U25 result to be more exposed. Arginine 52 interacts with U23 of the bulge and, in proximity to it, with G26 and A27 that therefore result to be fundamental for the Tat-TAR complex formation [39, 43].

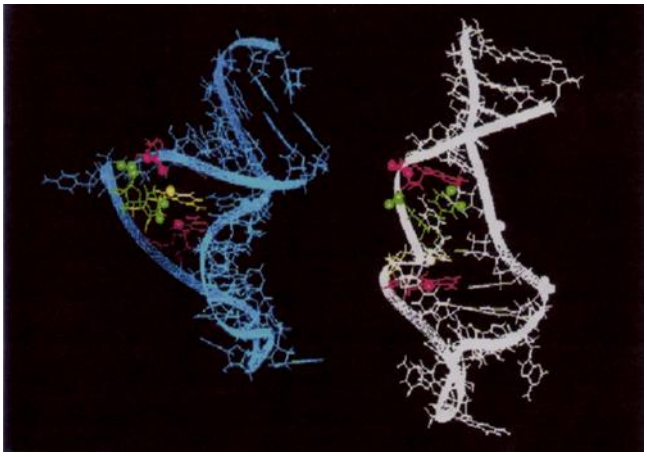


Figure 1.18 Representation of the major groove of free TAR (on the left) and of the Tat-bound TAR (on the right). A22 and A27 are represented in red; U23 in green and G26 in yellow [43].

1.1.13 Tat-TAR complex inhibition

The transactivation process represents a very attractive target for the development of potential anti-HIV agents. Small molecules that selectively bind the TAR-RNA viral sequence in order to prevent the Tat-TAR complexation can be divided into three categories of potential anti-HIV agents:

a. Peptide-based inhibitors;
b. Oligonucleotide-based inhibitors;
c. Non-peptide small molecules.

Peptide-based inhibitors. The first approach was based on the development of peptides similar to the Tat domain responsible of the binding to TAR, in order to obtain a selective interaction with the TAR bulge [44].

Oligonucleotide-based inhibitors. Oligonucleotides and oligonucleotide derivatives were studied as potential transactivation inhibitors due to their ability to form sequence-specific interaction with TAR. Antisense oligonucleotides complementary to the TAR stem-loop structure were studied and resulted to be active anti-HIV agents. An RNA aptamer against TAR, able to form a loop-loop complex, was in vitro selected and identified. Other oligonucleotides against TAR are RNA interfering (RNAi): RNAi are short double stranded

RNA sequences able to block HIV entry and replication through the blockade of cellular gene expression [45].

Non-peptide small molecules. In the last few years, small molecules interacting with TAR were deeply studied to inhibit the transcription of viral mRNA. These small molecules, as in literature reported, can interact with TAR in different way: small molecules can intercalate the TAR stem, probably in proximity to the bulge, or can interact directly with the bulge by stacking interactions [46].

Quinolones represent a promising class of small molecules studied as Tat-TAR complex inhibitors. Quinolones are antibacterial drugs that work by blocking the complex between gyrase enzyme and the bacterial chromosome; they have been hypothesized to be TAR binders and therefore inhibitors of Tat-TAR complex.

In the last few years, several studies reported the ability of quinolones to interfere with the Tat-mediated transcription process [47]. From SAR analysis performed on various derivatives, a class of ciprofloxacin derivatives that do not have the classical fluorine substituent in position 6 (6-desfluoroquinolones, 6-DFQs) resulted to be active and are now object of study to identify the chemical characteristic of the pharmacophore useful for obtaining more active analogues [48].

Structure–activity relationship (SAR) studies show that the highest antiretroviral activity was achieved when the essential 4-oxo-1,4-dihydroquinoline-3-carboxylic nucleus of these compounds was functionalized by a suitable 4-arylpiperazine substituent at the C7 position and an amino group or hydrogen atom at the C6 position. Although a methyl group emerged with optimal activity in the N1 position, the absence of a substituent at N1 also yielded acceptable anti-HIV activity. The 6-aminoquinolone WM5 is considered the prototype molecule from which to start the design of new antiviral derivatives (Figure 1.19).

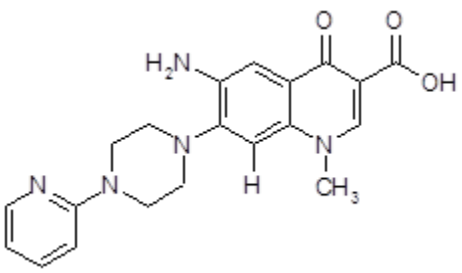


Figure 1.19. Chemical structure of quinolone WM5.

Focusing on the quinolone nucleus, new derivatives that have modifications at the N1 position were synthetized. The most relevant is the 1-amino derivative WP7-5 (Figure 1.20).

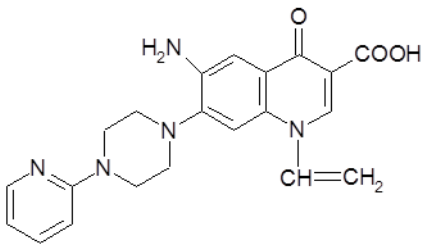


Figure 1.20 Chemical structure of WP7-5.

The ability of WP7-5 to inhibit the Tat-TAR complex formation was analyzed by FQA [49] and it resulted to be active as inhibitors with Ki value (0.89±0.09μM) lower than that of the 1-metil substituted WM5 (1.82±0.34μM) [50]. These results indicate that the N1-modification in the quinolone nucleus does not impair but rather improve the antiviral activity.

The synthesis of new WM5 analogues that have different chemical modifications at the N1 position could be useful to complete the screening of the 6-DFQs, in order to provide new insights into the SAR of antiviral quinolones.

1.2 AIM OF THE WORK

This research work is aimed to identify new anti-HIV agents having Reverse Transcription, a fundamental event in the replication cycle of HIV-1, as therapeutic target.

This thesis project had the following goals:

1. *Characterization of the full-length NC protein*

The first part of the work is aimed to deeply characterize the recombinant full-length NC protein and its biological activity. The analysis of the NC structure was completed by circular dichroic spectroscopy (CD) both in the presence and in the absence of zinc ions. Then, the chaperone activity of NC *i.e.* its ability to destabilize TAR and cTAR secondary structures was evaluated by a Fluorescence Assay and a FRET assay; NC aggregating activity was instead monitored through gel electrophoresis that allowed us to analyze the TAR/cTAR annealing reaction in the presence of the full-length NC protein.

2. *Identification of potential NC inhibitors*

The second goal was to perform a simple, fast and highly reproducible FRET-based assay for the High Throughput Screening of inhibitors of NC chaperone activity in order to identify and develop new potential antiviral agents.

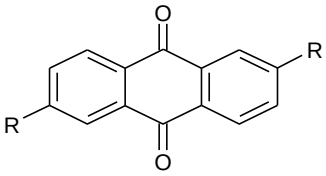
More of 100 small molecules from our in-house chemical library, representing a wide range of structures, were tested by HTS. In this thesis we focus on the investigation of 2,6-dipeptidyl anthraquinones, known DNA intercalators, which can also bind RNA sequences. It is reasonable to consider these anthraquinonic derivatives as potential NC inhibitors: they could interfere with NC chaperone activity thanks to their attitude of stabilize the secondary structures of RNA and DNA sequences involved in the retrotranscription process.

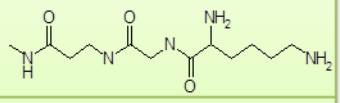
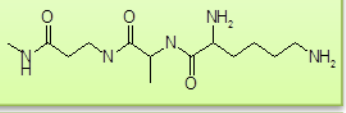
The proposed mechanism of action for anthraquionones was therefore studied in deep by Melting Assays (FRET probes coupled to thermal melting experiments). We analyzed the ability of each compound to bind, probably by intercalation, and stabilize the secondary structure of the viral nucleic acids cTAR, TAR and of the TAR/cTAR hybrid.

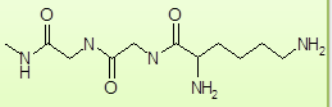
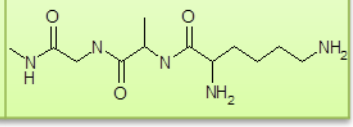
Finally, we analyzed the anthraquinones effect on the NC-mediated annealing of TAR with cTAR, to verify the effectiveness of the inhibition of the NC chaperon activity during the step subsequent to TAR and cTAR melting.

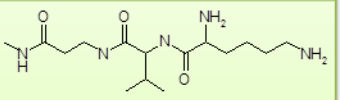
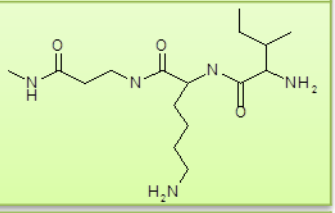
The chemical structures of tested 2,6-disubstitued anthraquinones, shown in Figure 1.21a (Z series), 1.21b (GSF series) and 1.21c (G series), contain an anthraquinonic nucleus able to intercalate into DNA and/or RNA and two identical side chains containing amino acid residues (with preference for basic ones).

Series Z



		R
Z6	2,6 AQ-βAla-Gly-Lys	
Z7	2,6 AQ-βAla-Ala-Lys	

		R
Z9	2,6 AQ-Gly-Gly-Lys	
Z8	2,6 AQ-Gly-Ala-Lys	

		R
542	2,6-AQ-βAla-Val-Lys	
543	2,6 AQ-βAla-Lys-Ile	

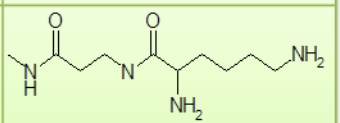
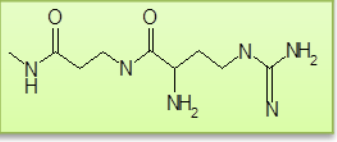
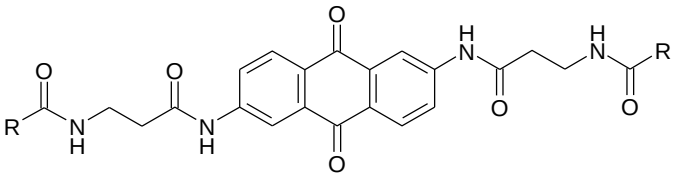
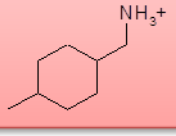
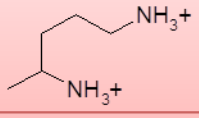
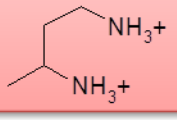
		R
721	2,6 AQ-βAla-Lys	
723	2,6 AQ-βAla-Arg	

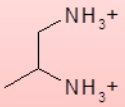
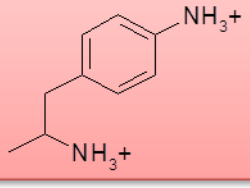
Figure 1.21a Chemical structures of 2,6-disubstituted anthraquinones of Z series, synthesized by Prof. G. Zagotto, Department of Pharmaceutical and Pharmacological Sciences, University of Padua.

Series GSF



	R
GSF1	
GSF3	

	R
GSF6	
GSF7	

	R
GSF8	
GSF11	

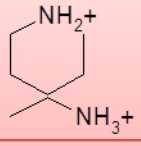
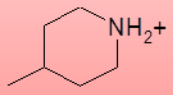
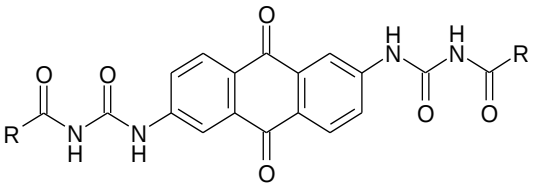
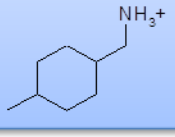
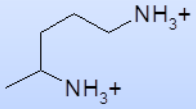
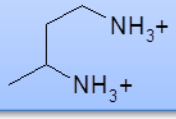
	R
GSF12	
GSF13	

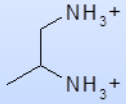
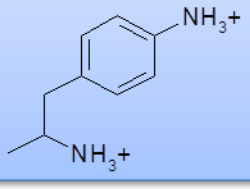
Figure 1.21b Chemical structures of 2,6-disubstitued anthraquinones of series GSF, synthesized by Prof. Santagada, Department of Pharmaceutical and Toxicological Chemistry, University of Napoli.

Series G



	R
G1	
G3	

	R
G6	
G7	

	R
G8	
G11	

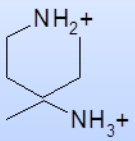
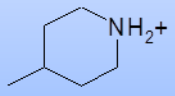
	R
G12	
G13	

Figure 1.21c Chemical structures of 2,6-disubstitued anthraquinones of series G, synthesized by Prof. Santagada, Department of Pharmaceutical and Toxicological Chemistry, University of Napoli.

The third goal was:

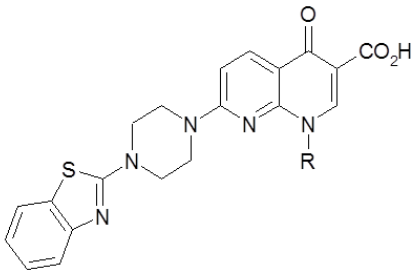
3. Identification of inhibitors of the Tat-mediated transcription process

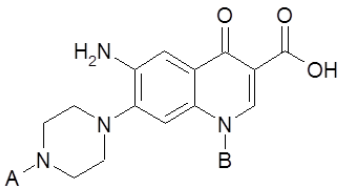
Two classes of compounds were studied as inhibitors of the Tat-TAR complex formation, a primary event of the transcription process.

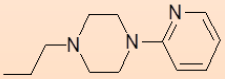
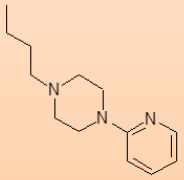
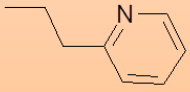
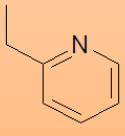
Quinolones are a promising class of anti-HIV compounds which interfere with the Tat mediated transcription process [47]. With the aim of provide new insights into the SAR of antiviral quinolones, WP and HP series were analyzed by Fluorescence Quanching Assay (FQA), a FRET-based platform of screening performed and validated in our lab [49, 50]. FQA allowed the determination of the inhibition constant Ki for each compound.

4-oxo-7-piperazin-1-yl naphtyridine-3-quinolinecarboxylic acid (WP series) and 4-oxo-7-piperazin-1-yl naphtyridine-3-carboxylic acid (HP series) were synthesized by Prof. Tabarrini, Department of Pharmaceutical Sciences, University of Perugia. The chemical structures of quinolones of WP and HP series are reported in figure 1.22.

HP and WP series





	R
HP6-13N	
HP3-13N	
HP12-13N	
HP10-13N	

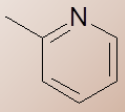
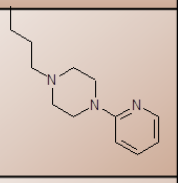
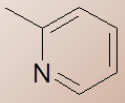
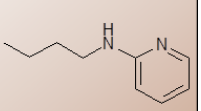
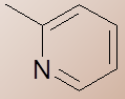
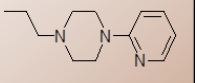
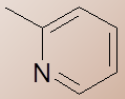
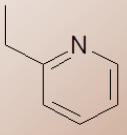
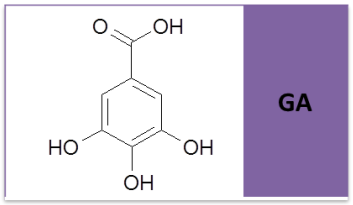
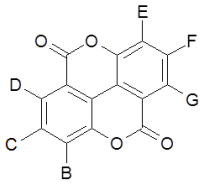
	A	B
WP 3-5		
WP 5-5		
WP 6-5		
WP 10-5		

Figure 1.22 Chemical structures of quionolones of HP and WP series.

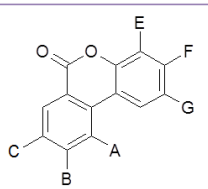
Ellagic acid (EA) and some synthetic bi- and mono-lactones derivatives proved to be promising active as chemopreventive agents and topoisomerases inhibitors [51]. Moreover, ellagic derivatives (ELs) showed also antiviral activities: some of them inhibit the replication of Herpes Simplex virus by blocking the viral adsorption to cultured cells and others inhibit the reverse transcriptase activity in mouse leukemia virus infected cells [52]. In this work ELs were analyzed to evaluate their inhibition activity on the Tat-TAR complex formation.
Ellagic acid (EA) and its derivatives were synthesized by Prof. G. Zagotto, Department of Pharmaceutical and Pharmacological Sciences, University of Padua, starting from EA, except for 1186 that was synthesized from two molecules of Gallic acid (GA). Their chemical structures are reported below (Figure 1.23).



GA



	B	C	D	E	F	G
EA	-OH	-OH	-H	-OH	-OH	-H
1181	-OCH3	-OCH3	-H	-OCH3	-OCH3	-H
1187	-OAc	-OAc	-H	-OAc	-OAc	-H
1186	-OH	-OH	-OH	-OH	-OH	-OH
1184	-SO4	-OH	-H	-SO4	-OH	-H



	A	B	C	E	F	G
1182	-OH	-OH	-OH	-OH	-OH	-H
1185	-OH	-OH	-OH	-OH	-OH	-OH
UA	-H	-H	-OH	-H	-OH	-H
UB	-H	-H	-H	-H	-OH	-H

Figure Table 1.23 Chemical structures of GA, EA and its bilactone natural and synthetic derivatives; chemical structures of Urolithins (monolactone derivatives) which are the main metabolites of the EA.

1.3 EXPERIMENTAL SECTION

1.3.1 Materials

Reagents

- Boric acid: H_3BO_3, PM=61.83, Sigma.
- APS (ammonium persulfate): $(NH_4)_2S_2O_8$, PM=228.20, Sigma.
- Hydrochloric acid: HCl, PM=36.46, Carlo Erba.
- Bromophenol blue: $C_{19}H_{10}BrO_4S$, PM=669.99, Sigma.
- Sodium chloride: NaCl, PM=58.44, Fluka.
- DMSO (Dimethyl sulfoxide): C_2H_6OS, PM=78.13, Riedel-de Haen.
- EDTA (Ethylenediaminetetraacetic acid dihydrate): $C_{10}H_{14}Na_2O_8$ x $2H_2O$, PM=372.24, Fluka.
- Ethanol: C_3H_5OH, PM=46.07, Fluka.
- Ethidium Bromide: GelRed^{TM} Nucleic Acid Gel Stain.
- Glycerol: $C_3H_8O_3$, PM=92.10, Fluka.
- Magnesium perchlorate: Mg(ClO4)2, PM=223.21, Sigma.
- TEMED (N,N,N',N' Tetramethylethylenediamine): $(CH_3)_2NCH_2CH_2N(CH_3)_2$, PM=116.2, Fluka.
- SDS (sodium dodecyl sulfate): $CH_3(CH_2)_{11}OSO_3Na$, PM=288.37, Sigma.
- Acrylamide (C_3H_5NO, PM=71.08)/N,N'-metilen-bis-acrylamide ($(CH_2CHCONH)_2CH_2$ PM=154.17) solution 19:1, Sigma.
- Sodium carbonate anhydrous: Na_2CO_3, PM=105.98, Fluka.
- Sodium thiosulfate pentahydrate: $Na_2S_2O_3$ x $5H_2O$, PM=248.18, Fluka.
- TRIS (trishydroxymethylaminomethane): $C_4H_{11}NO_3$, PM=121.10, Sigma.
- H_2O DEPC-treated (Fluka). DEPC (diethyl pyrocarbonate) is an efficient, nonspecific inhibitor of RNases and it reacts with amine, hydroxy and thiol groups of proteins thereby inactivating RNAses (and other enzymes). It is completely destroyed by autoclaving.
- MilliQ water.
- Coomassie Brillant Blue R250: $C_{45}H_{44}N_3NaO_7S_2$, PM=825.97, Sigma.

Buffers

- TN 1X (Tris HCl 10mM, NaCl 20mM, pH=7,5)
- TNMg 1X (Tris HCl 10mM, NaCl 20mM, $Mg(ClO_4)_2$ 1mM, pH=7,5)
- ETN 1X (Tris HCl 10mM, NaCl 20mM, EDTA 1 mM pH=7,5)
- TBE 1X (Tris 89mM, Boric acid 89 mM and EDTA 2 mM pH=8,0)

NC protein

The full-length NC protein (Figure 1.24), expressed and purified by Dr. Matteo Scalabrin (Dep. Chemistry, University at Albany), was aliquoted and conserved at -20°C. The protein concentration was controlled on a UV-Vis Spectrophotometer Lambda 20, Perkin Elmer, using an extinction coefficient at 280nm of 6410 M^{-1} cm^{-1}.

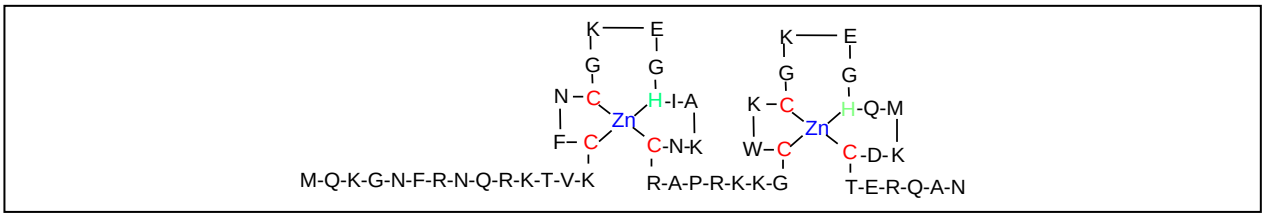


Figure 1.24 Sequence and secondary structure of the full-length HIV-1 NC protein, expressed by Dr. M. Scalabrin.

NC is characterized by two copies of a highly conserved retroviral-type, zinc-finger motif, also called CCHC motif that binds zinc ions strongly in a constrained structure. The chaperone activity of NC resides in the zinc-finger structures. Also the (12-55)NC peptide, containing the zinc-finger motifs but lacking the ability to aggregate the oligonucleotides was analyzed.

The (12-55)NC peptide was synthesized and purified by *EspIKem Peptides by design* (Polo scientifico e tecnologico di Sesto Fiorentino, Firenze, Italy). We used and analyzed different stocks of peptide: an aliquot of peptide referred as (12-55)NC was synthesized in the absence of zinc ions; another aliquot of peptide was synthesized in the presence of $ZnCl_2$ to preserve the highly oxidizable Cys residues and it was named (12-55)NCZn, since it was lyophilized in the zinc-bound form. The lyophilized peptides (12-55)NC and (12-55)NCZn were resuspended in H_2O or in TN 10/20 (Tris-HCl 10mM, NaCl 20mM pH 7.5).
The peptide concentration was determined on a UV-Vis Spectrophotometer Lambda 20, Perkin Elmer, using an extinction coefficient at 280nm of 5700 M^{-1} cm^{-1}.

Tat protein

The full-length Tat protein resulted to be unstable in solution, therefore all experiments were performed using the 10aa Tat peptide. The peptide sequence corresponds to the minimal amino acid sequence (48–57; GRKKRRQRRR) necessary for TAR binding. For the FQA the Tat-derived peptide labeled with the donor fluorescein at its N-terminus was used, while the unmodified 10aa Tat peptide was used for analysis by EMSA.

Nucleic acids

The oligonucleotides were purchased from Metabion International AG (Martinsried, Germany) and stored at −20 °C in TE (10mM Tris-HCl, 1mM EDTA) pH 8.0. Dilutions were all made in DEPC water. The sequences and their bulge-loop structures of TAR (RNA) and its complementary cTAR (DNA) are reported in Figure 1.25.

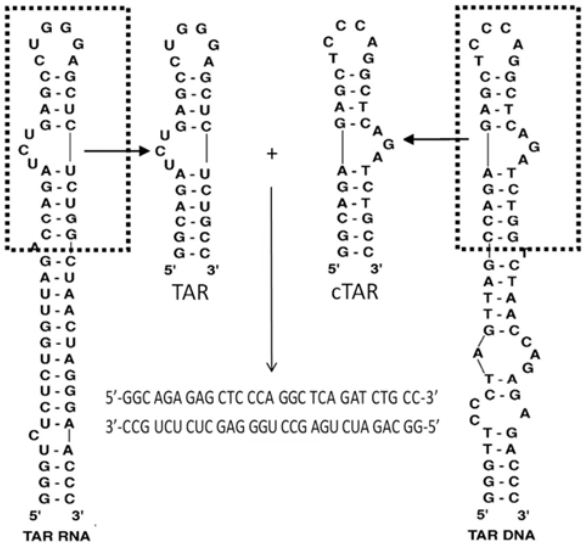
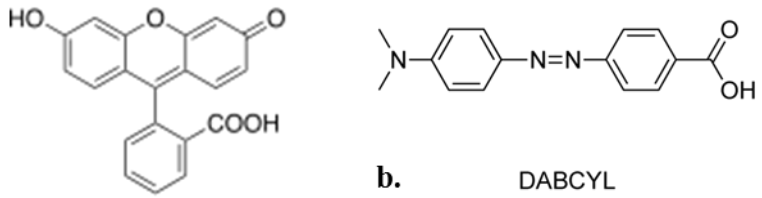


Figure 1.25 Sequence and structure of cTAR (29pb, PM=10183), TAR (29 bases, PM=10213,62) and TAR/cTAR hybrid.

For gel electrophoresis we used the unlabeled cTAR and TAR while for FRET assay, High Throughput Screening and Melting assay cTAR and TAR labeled at 5’ and 3’ ends respectively by the fluorophore 5-carboxyfluorescein (FAM) and the dark quencher 4-(4’-dimethylaminophenylazo)benzoic acid (DABCYL) were used. The 5’-dabcyl modified TAR was also synthetized to analyze the ELs interference on the Tat-TAR complex by FQA.
The chemical modifications introduced in TAR and cTAR are reported in Figure 1.26.



a. **b.** DABCYL

Figure 1.26 a) chemical structure of the fluorophore 6-carboxyfluorescein; b) chemical structure of the dark quencher DABCYL.

In order to promote the correct folding of TAR and cTAR (labeled or not) into a stem-loop structure, we used the following method: TAR (or cTAR) was diluted in TNMg 1X, denatured for 5 minutes at 100 °C and then left to cool to room temperature. This procedure allowed to renature correctly the oligonucleotides into a stem-loop structure corresponding to the wild type conformation. To promote the annealing of TAR with cTAR to form a DNA/RNA hybrid, we prepared a solution (in TNMg 1X) of the two oligonucleotides in equimolar concentrations and used the same protocol of denaturation and renaturation. This protocol is referred to as heat-annealing.

Compounds

2,6-dipeptidy anthraquinones of series Z, analyzed as NC inhibitors, and ELs, analyzed as inhibitors of the Tat-TAR complexation, were synthesized by Prof. G. Zagotto, Department of Pharmaceutical and Pharmacological Sciences, University of Padua.

Anthraquinones of series GSF and G, potential NC hinibitors, were synthesized by Prof. Santagada, Department of Pharmaceutical and Toxicological Chemistry, University of Napoli.

WP and HP quinolones series were synthesized by Prof. Tabarrini, Department of Pharmaceutical Sciences, University of Perugia.

The chemical structures of all tested compounds are reported in the chapter “Aim of the Work”.

1.3.2 Methods

CD analysis of the full-length NC structure

CD experiments were performed on a Jasco J-810 spectropolarimeter equipped with a Peltier temperature control system. Quartz cells with pathlength of 0.1 cm were used for the 197-250 nm regions. Each CD spectrum was an average of four accumulations scanned at 20 nm min^{-1}. Scans were obtained with bandwidth of 2 nm and response time of 16s. The protein concentration for far-UV CD was of 0.1 mg/ml in water. Observed ellipticities were converted to mean residue ellipticity $\theta = deg \times cm^2 \times dmol^{-1}$.

Fluorescence Assay of biological activity

To evaluate the biological activity of NC on TAR and cTAR a simple Fluoresce Assay was performed. 5'-FAM and 3'-DAB modified cTAR or TAR were used. 800µL of 0.1µM cTAR or TAR were folded in TNMg 1X (Tris-HCl 10mM, NaCl 20mM, $Mg(ClO_4)_2$ 1mM pH 7.5): the oligonucleotide was denatured at 95°C for 5 minutes and then left to cool to room temperature in order to assume the stem-loop structure, which is essential for the interaction with NC. The fluorescence signal of the folded cTAR or TAR was analyzed in cuvette with 494 nm as excitation and 500-700 nm as emission wavelengths. Increasing concentrations of NC were added to cTAR before reading. The fluorescence spectra were recorded also after adding an excess of EDTA, which removes Zn ions from Zn-fingers and finally after heating the sample.

FRET Assay of biological activity

The same experiment to evaluate the biological activity of NC on TAR and cTAR was performed using a microplate reader VictorIII (Perkin Elmer) which allows us to use lower volumes of sample, with 485 and 535nm as excitation and emission wavelengths. 5'-FAM and 3'-DAB modified cTAR or TAR (1uM) was folded in TNMg 1X (Tris-HCl 10mM, NaCl 20mM, $Mg(ClO_4)_2$ 0.2mM pH 7.5): the oligonucleotide was denatured at 95°C for 5 minutes and then left to cool to room temperature in order to assume the stem-loop structure. cTAR or TAR was then diluted to 0.1uM in TN 10/20 (Tris-HCl 10mM, NaCl 20mM pH 7.5). Increasing concentrations of NC (0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1 and 1.1uM) were added to 0.1uM cTAR or TAR in each well. The plate was read three times with a delay of 1 minute one reading from the other.

Analysis of the TAR/cTAR annealing reaction in the presence of the full-length NC

As controls, we used TAR, cTAR and the hybrid TAR/cTAR (each 1uM) folded in TNMg 1X (Tris-HCl 10mM, NaCl 20mM, Mg(ClO4)2 1mM pH 7.5): the oligonucleotides were denatured at 95°C for 5 minutes and then left to cool to room temperature in order to assume their stem-loop (TAR and cTAR) or double-stranded (hybrid TAR/cTAR) structure. The annealing reaction of the hybrid TAR/cTAR was performed both in the absence and in the presence of NC: TAR and cTAR, prefolded in TNMg 1X, were then incubated with NC (oligos:NC=1:8) for 3 hours at r.t. or at 37°C. The samples incubated with NC were added of native Gel Loading Buffer (Tris-HCl 10mM, EDTA 0.5mM, 50% w/v glycerol, 0.05% w/v bromophenol blue, 0.05% w/v xylene cyanol) or of denaturing Gel Loading Buffer (Tris-HCl 100mM, EDTA 4mM, 50% w/v glycerol, 2% w/v SDS, 0.05% w/v bromophenol blue). The samples were analyzed by electrophoresis on a 20% polyacrylamide gel at room temperature in TBE 1X (Tris-HCl 89mM, Borate 89mM, EDTA 2mM). After electrophoresis nucleic acids on the gel were stained with EtBr (0.5ug/mL, 30' under agitation); nucleic acids fluorescence in the gel systems was detected on a Geliance 600 Imaging System (PerkinElmer).

High Throughput Screening (HTS) by FRET Assay

The high throughput screening of potential inhibitors of the chaperone activity of NC was performed by FRET Assay, using a microplate reader VictorIII (Perkin Elmer) with 485 and 535nm as excitation and emission wavelengths. 5'-FAM and 3'-DAB modified cTAR or TAR (1µM) was folded in TNMg 1X (Tris-HCl 10mM, NaCl 20mM, Mg(ClO4)2 0.2mM pH 7.5): the oligonucleotide was denatured at 95°C for 5 minutes and then left to cool to room temperature in order to assume the stem-loop structure. cTAR or TAR was then diluted to 0.1uM in TN 10/20 (Tris-HCl 10mM, NaCl 20mM pH 7.5). Increasing concentrations of compound (0, 0.1, 0.5, 1, 5, 10, 50, 100uM) NC were incubated with 0.1µM cTAR or TAR in each well. Finally, NC 0.8uM (molar ratio oligo/NC=1/8) was added to each sample. The plate was read three times with a delay of 1 minute one reading from the other. The experimental data were fitted and the IC_{50} value was calculated for each compound.

Each experiment was performed in triplicate to calculate a standard deviation of the IC_{50} value. IC_{50} is defined as the measure of the compound effectiveness in inhibiting biological or biochemical function. This quantitative measure indicates how much of a particular drug or substance is needed to inhibit a given biological process by half. It is commonly used as a measure of drug potency in pharmacological research. The IC_{50} of a drug can be determined

by constructing a dose-response curve and examining the effect of different concentrations of compound, on biological process. In general, the higher the concentration of drug, the more the biological process can be reduced.

IC_{50} values can be used to compare the potency of different tested compounds. IC_{50} values are dependent on experimental conditions, for instance on substrate concentration: if the concentration of substrate increases, the value of IC_{50} raises.

Data were plotted on a graph reporting in abscissa the compound concentrations in logarithmic scale (Log [Molar concentration] μM) and in ordinate the fractional activity expressed as:

$$Y = {(F_x - F_o)}\Big/{(F_m - F_o)}$$

Where F_x is the oligonucleotide fluorescence in the presence of NC and of the compound at a certain concentration. While F_o and F_m are the nucleic acid fluorescence in the absence of NC (negative control) and the reference value for the 100% activity respectively.

The resulting curve was a sigmoid similar to the following one in Figure 1.27:

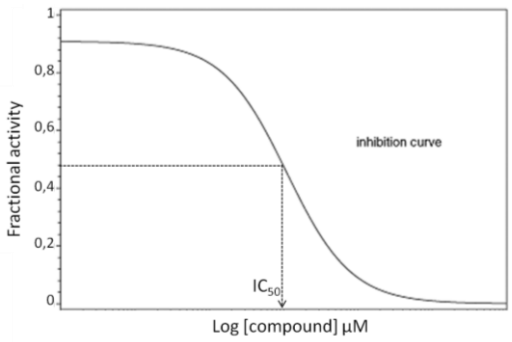


Figure 1.27 Sigmoid dose-response curve used to determine the IC_{50} values

To calculate the IC_{50} values, the experimental data were fitted using GraphPad Prism 5 software with the following equation:

$$Y = Ymin + \frac{(Ymax - Ymin)}{(1 + 10^{(x - LogIC50)})}$$

$$IC_{50} = \frac{10^x (Y - Ymin)}{(Ymax - Y)}$$

Where $Ymin$ and $Ymax$ are the lowest and highest value of fractional activity obtained respectively. While x and Y are the concentration of the tested compound (expressed in

logarithmic scale, Log[M] μM) and the value of fractional activity for that sample concentration respectively.

Melting Assay

The ability of each compound to interact with viral sequences was analysed by Melting Assay. The melting temperature (*Tm*) of the oligonucleotides, in the presence and in the absence of the compound, was measured with the purpose to calculate the variation of this value (ΔTm). Melting Assay allows the evaluation of compounds capability to alter nucleic acids thermal denaturing profile, due to direct interaction and structure stabilization. This assay was useful to understand the kind of interaction that subsists between compound and the nucleic acids.

The calculation of a positive ΔTm value means that the compound interacts with nucleic acid, probably by intercalation in the case of anthraquinones, restraining melting. When such a molecule is added to a oligonucleotide sequence, is able to stabilize nucleic acid conformation. The increment in Tm is proportional to compound concentration. While if the ΔTm value turns out to be equal to zero it suggests that the tested molecule does not interact with nucleic acids. If the ΔTm value results to be negative it means that the tested molecule destabilizes nucleic acids secondary structure and facilitates their thermal denaturation. The aim of the experiment was to understand if the impairment of NC-induced cTAR (or TAR) melting was the result of the compound interaction with nucleic acids.

Melting Assay was performed using 5'-FAM and 3'-DAB modified cTAR and TAR. TAR and cTAR. TAR or cTAR was folded in TNMg 1X at a concentration of 10μM and then diluted in ETN 1X buffer. The same protocol was applied for DNA/RNA hybrid: doubly labeled cTAR and TAR, both at 10μM concentration, were annealed together in TNMg 1X and then diluted in ETN 1X buffer. In each microplate well the oligonucleotide solutions (1μM) of TAR and cTAR were mixed with the sample solution (1, 10 and 100μM). The final volume in each well was 20μl. In every microplate control wells were: cTAR (or TAR, or TAR/cTAR hybrid) solution without compound (5 μl of MilliQ water) as reference value for *Tm* and the oligonucleotide solution in the presence of the highest concentration of DMSO tested. Melting Assay was performed twice for each compound-oligonucleotide pairs. The temperature was increased from 25 to 99°C in 1h (0,02°C/s) to promote the thermal denaturation of the oligonucleotides while the emission fluorescence (at 510nm) was monitored (30 acquisitions per °C).

The *Tm* value was mathematically derived from the thermal denaturing profile using LC480 software. *ΔTm* was calculated using the following equation:

$$\Delta Tm = Tm._{2} - Tm._{1}$$

Where $Tm._{2}$ and $Tm._{1}$ are the Tm values measured testing the oligonucleotides (or the hybrid) in the presence and in the absence of compound respectively.

Inhibition of the TAR/cTAR annealing reaction by anthraquinones

As controls, we used TAR, cTAR and the hybrid TAR/cTAR (each 1uM) folded in TNMg 1X (Tris-HCl 10mM, NaCl 20mM, Mg(ClO4)2 1mM pH 7.5): the oligonucleotides were denatured at 95°C for 5 minutes and then left to cool to room temperature in order to assume their stem-loop (TAR and cTAR) or double-stranded (hybrid TAR/cTAR) structure. TAR and cTAR were incubated separately with increasing concentration of compound (0, 1, 10, 50, 100uM) for 15 minutes at room temperature. TAR and cTAR pre-incubated with the compound were then mixed and incubated with NC 8uM (oligos:NC=1:8) for 15 minutes at room temperature. The samples were added of denaturing Gel Loading Buffer (Tris-HCl 100mM, EDTA 4mM, 50% w/v glycerol, 2% w/v SDS, 0.05% w/v bromophenol blue). The samples were kept steady on ice until they were loaded into the gel. The samples were analyzed by electrophoresis on a 12% polyacrylamide gel at room temperature in TBE 1X (Tris-HCl 89mM, Borate 89mM, EDTA 2mM). After electrophoresis nucleic acids on the gel were stained with SybrGreen; nucleic acids fluorescence in the gel systems was detected on a Geliance 600 Imaging System (PerkinElmer).

Fluorescence Quenching Assay (FQA)

The effect of quinolones of WP and HP series on the Tat/TAR complex was evaluated using FQA, a FRET-based competition assay performed using the Tat-derived peptide labeled with the donor fluorescein at its N-terminus and the 29 nt TAR labeled at its 3’ end with the dark quencher dabcyl. FQA was performed and validated in our lab in the last few years [Giaretta].

The peptide sequence corresponds to the minimal amino acid sequence (48–57; GRKKRRQRRR) necessary for TAR binding (5’-GGC AGA UCU GAG CCU GGG AGC UCU CUG CC-3’). Quenching of the Tat peptide fluorescence occurs upon incubation with increasing concentration of the quenching nucleic acid; disruption of peptide–RNA binding

by competing quinolones is observed as an increase in donor fluorescence intensity. Titrations of labeled Tat/TAR in the presence of competitors were performed in triplicate in a 96-well plate reader (Victor III, Perkin–Elmer); the fluorescein-labeled peptide was excited at 490 nm and emission was recorded at 535 nm. The plates were assembled with 190 mL of a solution containing 10 nM fluorescein–peptide in 10 mM Tris, 1 mM $Mg(ClO_4)_2$, 20mM NaCl, and 0.01% Triton X-100 (for the ten-amino acid peptide), in the presence of 10 µM quinolone. Fluorescence intensities in the presence of increasing concentrations of dabcyl-TAR were measured and curves were fitted to obtain Ki values. Preliminary determination of the spectral parameters of all quinolones allowed us to exclude optical interferences with fluorescein–peptide signal within the conditions of the assay.

EMSA (Electrophoresis Mobility Shift Assay)

ELs analysis was not possible to be performed by FQA due to a quenching effect observing after the addition of ELs to fluorescein-labeled Tat. The analysis of the ELs interference on the Tat-TAR complex was therefore performed by EMSA.

Unmodified Tat 10aa (10µM) was incubated with increasing concentrations of each compound (0, 1, 5, 10, 50, 100µM final) at room temperature for 15 minutes and then 60ng of folded TAR was added for other 15’ minutes. The samples were resolved by 10% non-denaturing polyacrylamide gels containing 1X TBE buffer. After electrophoresis nucleic acids on the gel were stained with SybrGreenII; nucleic acids fluorescence in the gel were detected on a Geliance 600 Imaging System (PerkinElmer).

1.4 RESULTS AND DISCUSSION

1.4.1 Full-length NC protein: structure and biological activity

1.4.1.1 Full-length NC protein: structure

The nucleocapsid protein NCp7 of human immunodeficiency virus type 1 (HIV-1) is a small, basic protein that is tightly associated with genomic RNA in the mature infectious virus. The nucleic acid aggregating activity of NC resides primarily in the basic N-terminal domain and the duplex destabilizing activity of NC has been mapped to its zinc finger structures [53, 54].

Many works in literature suggest that, to assay the chaperone-like NC activity, (12-55)NC peptide, that corresponds to the Zn fingers domain of NC, is preferred to the native NC because it preserves the nucleic acid binding and chaperone properties of the native NC [24, 33] but does not aggregate the oligonucleotides [19]. For this reason, our original project aimed to develop a platform to screen anti-NC small molecules employing the (12-55)NC peptide and so we began to use and analyze different stocks of peptide. The first aliquot of peptide referred as (12-55)NC was synthesized in the absence of zinc ions; the second aliquot of peptide was synthesized in the presence of $ZnCl_2$ to preserve the highly oxidizable Cys residues and it was named (12-55)NCZn, since it was lyophilized in the zinc-bound form. The lyophilized peptides (12-55)NC and (12-55)NCZn were resuspended in H_2O or in TN buffer. Both aliquots were analyzed: the investigation of the structure and of the biological functionality is reported in Appendix: the results obtained were not reproducible and not useful to achieve our goal. The use of the (12-55)NC peptide was abandoned and we focused on the recombinant full-length NC protein.

The full-length NC protein used in this work was expressed and purified by Fabris group at Albany: the protein identity was confirmed by MS analysis and the protein concentration was controlled by UV-Vis spectra (Scalabrin M., data not shown). The characterization of NC structure was completed by circular dichroic spectroscopy both in the presence and absence of zinc ions (Figure 1.28).

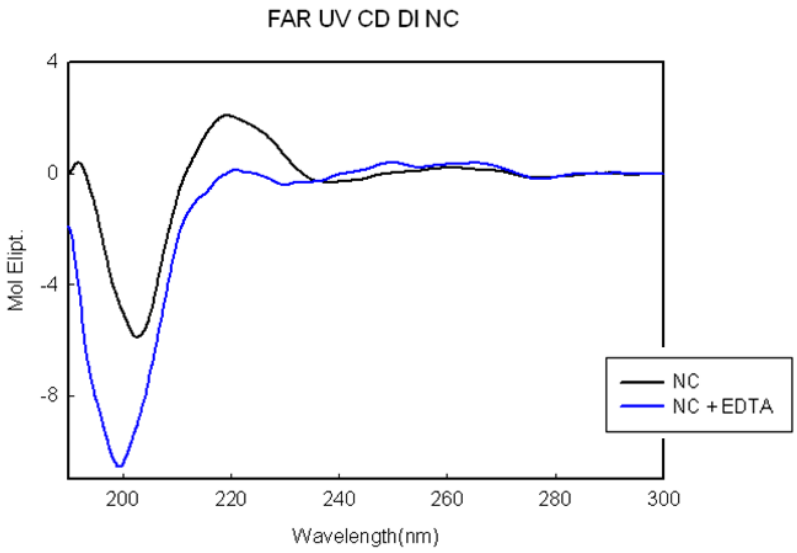


Figure 1.28 Circular dichroism (CD) spectra of the full-length NC protein in the presence (black line) and in the absence (blu line) of zinc ions.

CD spectra were recorded from 190 to 260 nm at 25°C. The CD spectrum of the full-length protein analyzed in the Zn-bound structure is reported in Figure 1.28 (black line): the spectrum is characterized by a broad band with a positive maximum at ≈219nm and a stronger band with negative ellipticity centered at ≈202nm, indicative of Zn finger structures [55, 56]. EDTA was then added to the sample. EDTA is a chelating agent that can remove Zn ions from the Zn-finger motifs. The spectrum recorded is reported in Figure 1.28 (blue line). In the absence of Zn ions, there is a significant increase in the intensity of the negative band as well as a slight blue shift (minimum at 198 nm), the positive band also suffers a decrease in intensity: the spectrum is indicative of a random coil [55, 56].

1.4.1.2 Full-length NC protein: biological activity

Fluorescence Assay of biological activity

NC possesses a nucleic acid chaperone activity that is critical in minus- and plus-strand transfers during reverse transcription. The minus strand transfer notably relies on the ability of NC to destabilize the stable stem with five contiguous, double stranded segments of both the TAR sequence of the viral genome and the complementary sequence, cTAR [57].

A simple fluorescence assay was performed to evaluate the chaperone activity of NC *i.e.* the ability of NC to destabilize the lower stem of the bulge-loop structured TAR and cTAR. The fluorescence spectra of the analysis of the full-length NC biological activity on cTAR are reported in Figure 1.29a and on TAR are reported in Figure 1.29b.

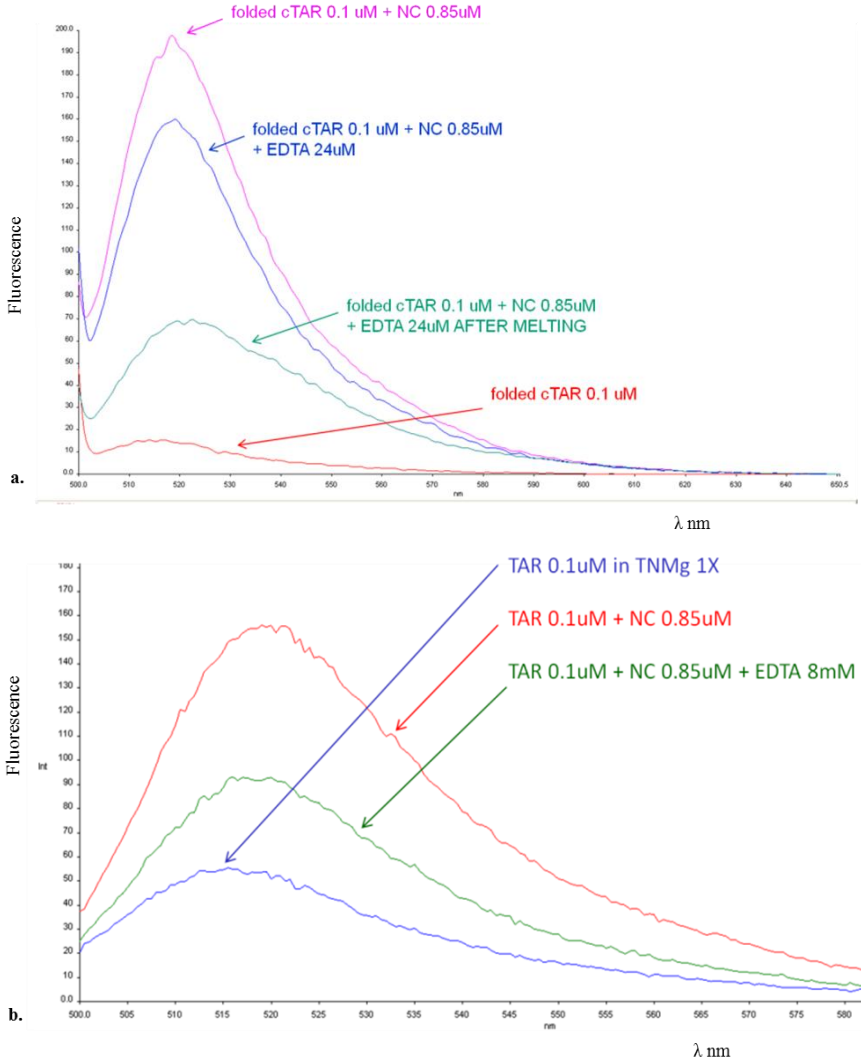


Figure 1.29 a) Fluorescence Assay: analysis of the biological activity of the full-length NC protein on cTAR. Fluorescence spectra of 0.1μM 3'-DAB and 5'-FAM modified cTAR in TNMg (Tris 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM, pH 7.5) in absence (red line) and in presence (pink line) of NC; after adding an excess of EDTA (blu line), which removes Zn ions from Zn-fingers and finally after heating the sample (green line). **b) Fluorescence Assay: analysis of the biological activity of the full-length NC protein on TAR.** Fluorescence

spectra of 0.1μM 3'-DAB and 5'-FAM modified TAR in TNMg 1X (Tris 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM, pH 7.5) in absence (blue line) and in presence (red line) of NC; green line represents the spectrum after adding an excess of EDTA which removes Zn ions from Zn-fingers. Spectra were corrected for buffer fluorescence.

In the case of cTAR, the addition of NC determines a rise in fluorescence , due to the increase of the distance between DAB and FAM that accompanies the destabilization of the cTAR stem-structure. Therefore the NC protein can act as chaperone in these conditions. Moreover, we can observe the effect of the addition of an excess of EDTA: the addition of EDTA determines a fluorescence decrease. EDTA is a chelating agent that can remove Zn ions from the Zn-finger motif and reduce the protein activity. If the protein is not correctly folded, it cannot act as chaperone on cTAR. After heating the sample, the system collapses and the fluorescence decreases. The NC protein can act as chaperone also on TAR in these conditions (Figure 1.29b). Hence, the protein is active as nucleic acids chaperone on both TAR and cTAR.

After having tested the chaperone activity of the Zn-bound NC on TAR, the fluorescence assay on TAR was further performed using this Zn-unbound structured NC, *i.e.* the NC protein without the Zn finger structures due to the ejection of Zn ions by EDTA. The added NC in this case was not correctly folded (confirmed by CD spectra, Figure 1.28). The fluorescence spectra are reported in Figure 1.30.

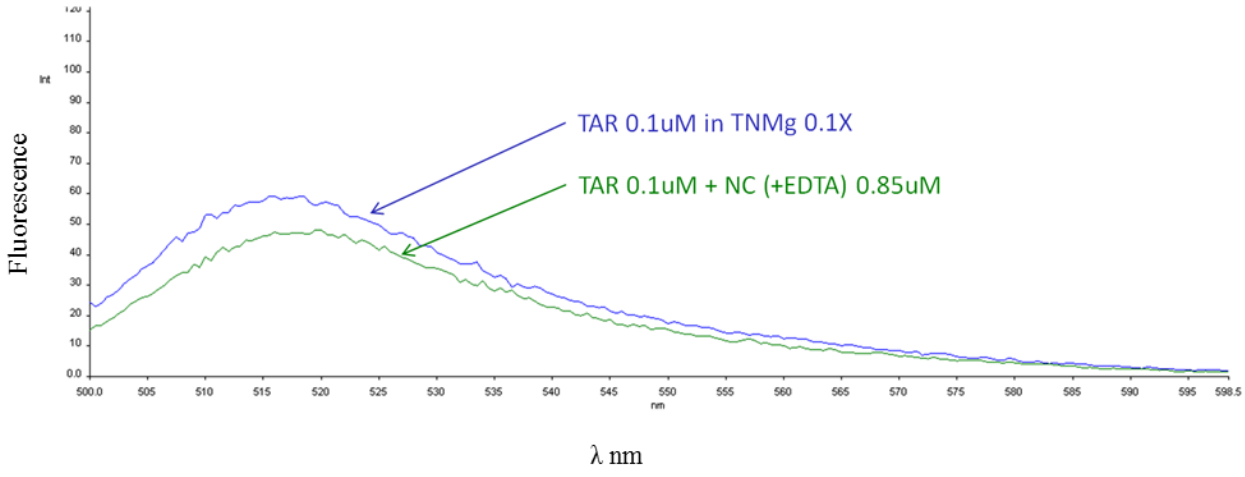


Figure 1.30 Fluorescence Assay: analysis of the biological activity of the Zn-unbound structured NC protein on TAR. Fluorescence spectra of 0.1μM 3'-DAB and 5'-FAM modified TAR in TNMg 1X (Tris 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM, pH 7.5) in absence (blue line) and in presence (green line) of NC. Spectra were corrected for buffer fluorescence.

In this case, the addition of NC does not determine an increase of fluorescence. The nucleic acids destabilizing activity of NC has been mapped to its zinc finger structures [53,

54], therefore if the protein is not correctly folded, it cannot act as chaperone on TAR. We can also observe that the presence in solution of Zn ions, ejected by EDTA, determines a quenching effect on the fluorescence signal, according to further results reported below.

Literature data suggest that divalent metal ions are potent fluorescence quenchers, therefore we decided to evaluate the effect of divalent Zn ions in solution on the fluorescence signal of fluorescein that is the fluorophore conjugated to TAR and cTAR in our Fluorescence Assay. The fluorescence signal of free fluorescein is very high compared with the FAM-DNA conjugate; the analysis was therefore performed on a 0.01uM fluorescein solution in TNMg 1X. The extent of quenching is dependent not only on the type of divalent cations but also on the presence of other ions in solution [58] and therefore EDTA was added and analyzed in the system. The spectra are reported in Figure 1.31.

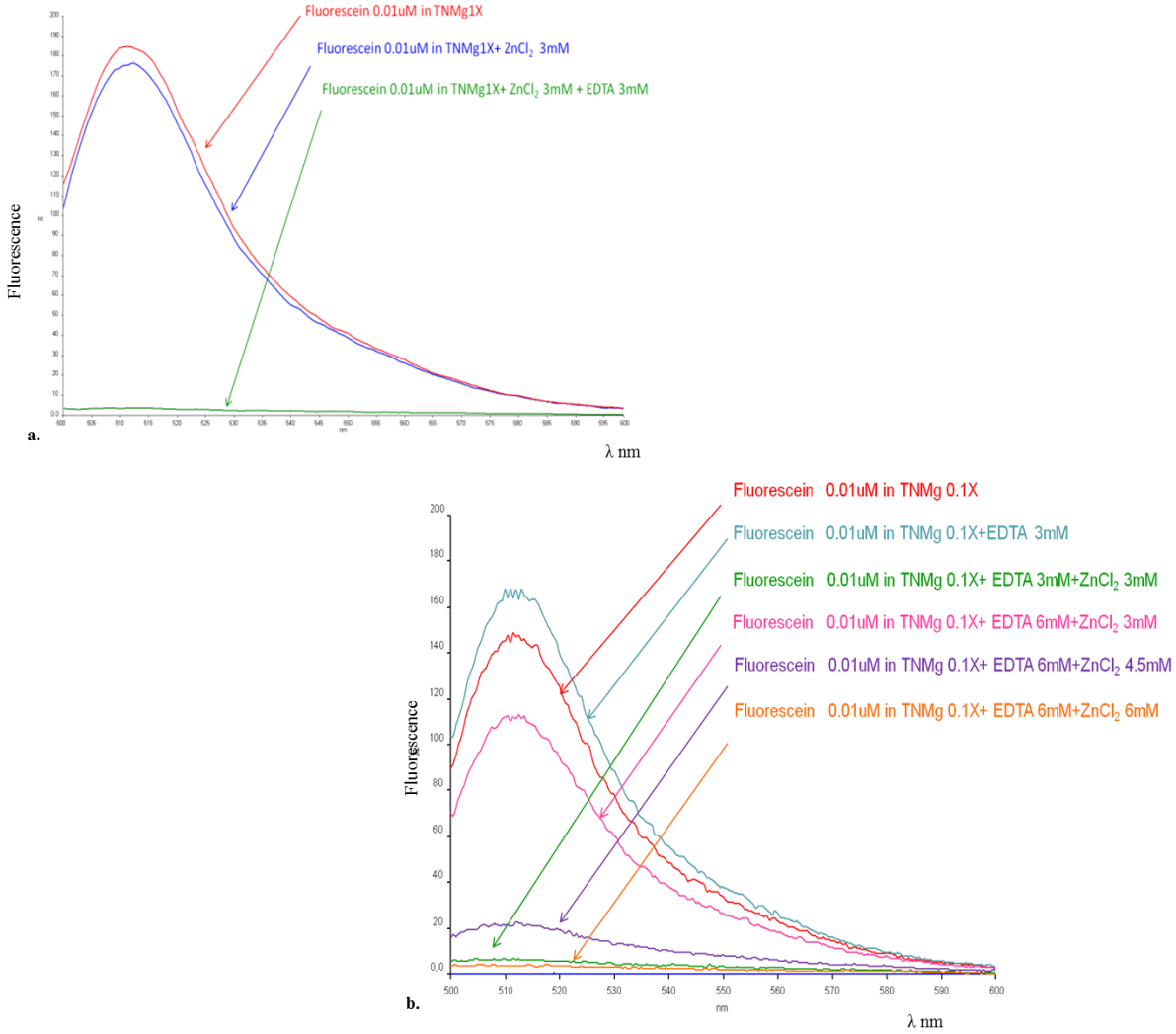


Figure 1.31 Fluorescence Analysis: effect of EDTA and Zn ions on fluorescein fluorescence. Fluorescence spectra of 0.01µM fluorescein in TNMg 1X (Tris 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM, pH 7.5). **a.** $ZnCl_2$ 3mM (blu line) and then EDTA 3mM (green line) were added to fluorescein in TNMg 1X (red line). **b.**

Fluorescence spectra of Fluorescein 0.01uM in TNMg 1X (red line) with EDTA 3mM (light blue line), EDTA 3mM and $ZnCl_2$ 3mM (green line), EDTA 6mM and $ZnCl_2$ 3mM (pink line), EDTA 6mM and $ZnCl_2$ 4.5mM (violet line), EDTA 6mM and $ZnCl_2$ 6mM (orange line). Spectra were corrected for buffer fluorescence.

In figure 1.31a, we can observe that the addition of $ZnCl_2$ determines a decrease of fluorescein fluorescence (blue line), but the simultaneous presence of Zn-EDTA determines a complete fluorescence quenching (green line). On the contrary, the addition of 3mM EDTA to fluorescein determines a moderate increase of fluorescence intensity (Figure 1.31b, light blue line), but the following addition of an equimolar amount of $ZnCl_2$ (3mM) results in the complete fluorescence quenching (green line). To follow out the effect of EDTA, Zn ions and of the EDTA-Zn complex on fluorescence, we continued the analysis with adding an excess of EDTA (6mM final) which determines the fluorescence increase, then the addition of $ZnCl_2$ (4.5mM final) causes a fluorescence quenching and finally the addition of an equimolar amount of $ZnCl_2$ (6mM final) results in the complete fluorescence quenching (Figure 1.31b).

Hence, the effect of EDTA on fluorescence is a moderate increase of fluorescence; Zn ions in solution determine a decrease of fluorescence and the simultaneous presence of EDTA and $ZnCl_2$ *i.e.* Zn-EDTA complex results in the complete fluorescence quenching of the free fluorescein.

FRET Assay of biological activity

The fluorescence assay to evaluate the biological activity of NC was then carried out using the microplate reader VictorIII which allows us to use lower volumes of sample. In figure 1.32 are reported the results of the FRET Assay, *i.e.* the variations of fluorescence induced on cTAR (Figure 1.32a) and on TAR (Figure 1.32b) by increasing concentrations of protein.

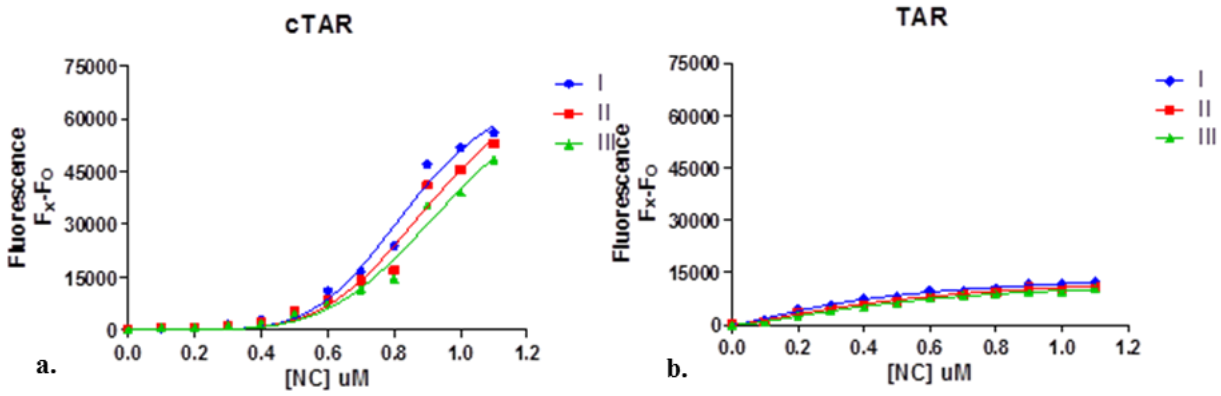


Figure 1.32 a. Fluorescence Assay: analysis of the biological activity of the full-length NC protein on cTAR. Fluorescence emission of 0.1μM 3'-DAB and 5'-FAM modified cTAR in presence of increasing concentrations of NC. The plate was read three times with a delay of 1 minute one reading from the other. **b.**

Fluorescence Assay: analysis of the biological activity of the full-length NC protein on TAR . Fluorescence emission of 0.1μM 3'-DAB and 5'-FAM modified TAR in presence of increasing concentrations of NC. The plate was read three times with a delay of 1 minute one reading from the other.

The recombinant protein results active on both the bulge-loop structures cTAR (Figure 1.32a) and TAR (figure 1.32b): the addition of NC determines the destabilization of the stem and the consequent increase of the distance between DAB and FAM that results in the rise of fluorescence. The better compromise solution to have the higher chaperone activity and the most stability over time for both cTAR and TAR is the ratio nucleic acid/NC=1/8. This result is in agreement with literature data: the cTAR sequence binds eight NC molecules at saturating protein concentration [24, 27].

1.4.1.3 Analysis of the TAR/cTAR annealing reaction in the presence of the full-length NC

NC possesses several activities in addition to the melting of nucleic acids stem-loop structures described above: one additional activity is the annealing between nucleic acid strands, reducing their inter-strand repulsion. The protein is positively charged and lowers the electrostatic barrier of the annealing reaction and increases the rate at which two separate complementary sequences comes together. Unstructured oligonucleotides usually anneals rapidly, while for structured nucleic acids a chaperone protein is functional to facilitate their association, after having reorganized their secondary structures [59]. This competence is connected with the chaperone activity and with the ability of the protein to aggregate nucleic acids (due to its unstructured N-terminal region). The analysis of the TAR/cTAR annealing reaction in the presence of NC, under conditions previously identified by fluorescence assays, was investigated by electrophoresis.

As controls, we used TAR, cTAR and the hybrid TAR/cTAR. The annealing reaction of the hybrid TAR/cTAR was performed both in the absence and in the presence of NC: TAR and cTAR were prefolded in TNMg 1X and then NC was incubated a few minutes at room temperature with the nucleic acids (oligos:NC=1:8). In the experiment reported in Figure 1.33, each sample was added of native Gel Loading Buffer.

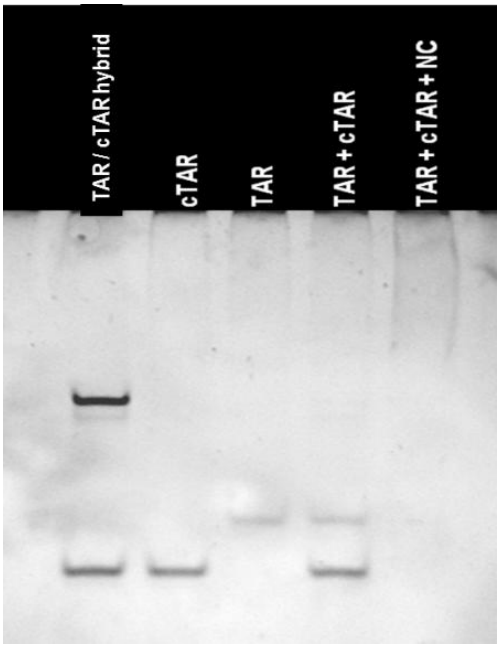


Figure 1.33 Analysis of the TAR/cTAR annealing reaction. Controls: TAR, cTAR and the hybrid TAR/cTAR (each 1uM) folded in TNMg 1X; annealing of TAR to cTAR in the absence and in the presence of NC. Each sample was added of native Gel Loading Buffer. Electrophoresis on a 20% polyacrylamide gel at r.t. in TBE 1X. After electrophoresis nucleic acids on the gel were stained with EtBr; nucleic acids fluorescence in the gel systems was detected on a Geliance 600 Imaging System.

The annealing of the TAR/cTAR hybrid by heating was verified. The annealing reaction of prefolded TAR to prefolded cTAR was performed both in the absence and in the presence of NC: without NC, TAR and cTAR move like their controls; in the presence of NC, we cannot detect the oligos but a smear up to the well, indicating aggregation by NC: protein-nucleic acids complexes, hence, cannot migrate into the gel. Therefore, we can affirm that TAR and cTAR interact with NC, but we cannot detect the hybrid formation. Moreover, we can observe that EtBr staining is different for DNA and RNA oligonucleotides, therefore the evaluation of the reaction can be qualitative and not quantitative.

To improve the annealing reaction, the assay was performed introducing some modifications. The controls used were the stem-loop structured TAR, cTAR and the double-stranded hybrid TAR/cTAR folded in TNMg 1X by heating. The annealing reaction of TAR to cTAR was performed both in the absence and in the presence of NC: TAR and cTAR, prefolded in TNMg 1X, were then incubated with NC (oligos:NC=1:8) for 3 hours at r.t. or at 37°C. The samples incubated with NC were added of native Gel Loading Buffer or of denaturing Gel Loading Buffer. In Figure 1.34 is reported the image of the gel.

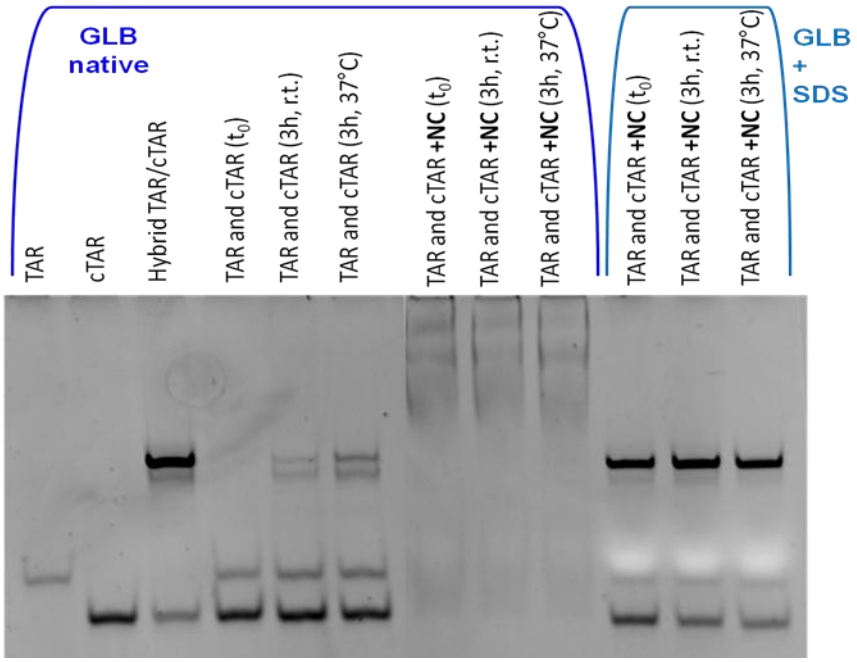


Figure 1.34 Analysis of the TAR/cTAR annealing reaction in the presence of the full-length NC protein. Controls: TAR, cTAR and the hybrid TAR/cTAR (each 1uM) folded in TNMg 1X; annealing of TAR to cTAR in the absence and in the presence of NC: TAR and cTAR, prefolded in TNMg 1X, were incubated with NC (oligos:NC=1:8) for 3 hours at r.t. or at 37°C. The samples incubated with NC were added of native Gel Loading Buffer or of denaturing Gel Loading Buffer. Electrophoresis on a 20% polyacrylamide gel at r.t. in TBE 1X. After electrophoresis nucleic acids on the gel were stained with EtBr; nucleic acids fluorescence in the gel were detected on a Geliance 600 Imaging System (PerkinElmer).

The annealing of the hybrid TAR/cTAR by heating was verified. We can compare the annealing reaction of TAR to cTAR in the absence and in the presence of NC: without NC the hybrid TAR/cTAR was formed in a low percentage after 3 hours both at r.t. and at 37°C; in the presence of NC, we cannot detect the oligos when the samples were added of native GLB because they are aggregated to NC, but we can detect the hybrid formation when the samples were added of denaturing GLB. In the presence of the denaturing GLB containing 2% w/v SDS, NC results denatured and the oligos, protein-free, can migrate into the gel.

In the absence of NC the annealing reaction is extremely slow; on the other hand, the hybrid TAR/cTAR formation reaction is dramatically accelerated by NC. Moreover, we can observe that without NC there are two different bands with similar electrophoretic mobility corresponding to the hybrid, while in the presence of NC the band corresponding to TAR/cTAR prevails over the other band. These results confirm that in the absence of NC the annealing mechanism involves the formation of a 'kissing complex' formed by the apical loops of TAR and cTAR; in the presence of NC, the pathway of the annealing reaction is clearly shifted toward the zipper mode [60].

The annealing reaction NC mediated was further characterized: NC was titrated respect to 1uM of both TAR and cTAR (Figure 1.35a) to identify the lower NC concentration necessary to obtain the complete hybrid TAR/cTAR formation and then a time drive analysis was performed (Figure 1.35b).

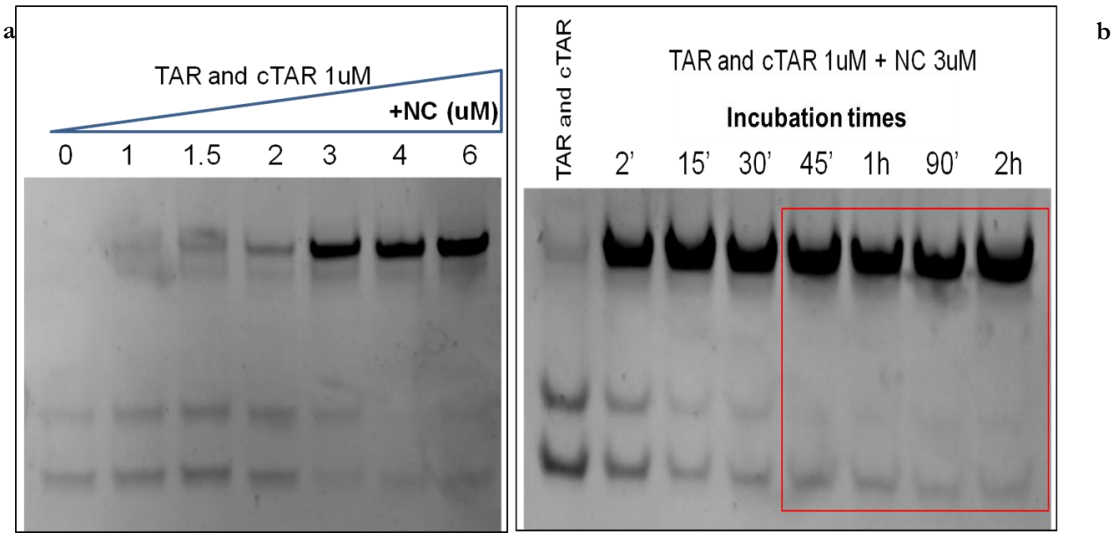


Figure 1.35 a. NC titration: increasing concentration of NC (0, 1, 1.5, 2, 3, 4, 6uM) were added at r.t. for 10 minutes to TAR and cTAR (each 1uM) prefolded in TNMg 1X. **b. Kinetics of the annealing reaction**: 3uM of NC was added to TAR and cTAR (each 1uM) prefolded in TNMg 1X at r.t. for 2, 15, 30, 45, 60,90, 120 minutes. The samples were added of denaturing Gel Loading Buffer (Tris-HCl 100mM, EDTA 4mM, 50% w/v glycerol, 2% w/v SDS, 0.05% w/v bromophenol blue). All samples were analyzed by electrophoresis on a 20% polyacrylamide gel at room temperature in TBE 1X. After electrophoresis nucleic acids on the gel were stained with EtBr (0.5ug/mL, 30' under agitation); nucleic acids fluorescence in the gel were detected on a Geliance 600 Imaging System (PerkinElmer).

The annealing reaction of TAR to cTAR was performed in the presence of increasing concentrations of NC (0, 1, 1.5, 2, 3, 4, 6uM). Figure 1.35a shows that the lower NC concentration necessary to obtain the complete hybrid TAR/cTAR formation is 3uM (oligos:NC=1:3).

Then, the kinetics of the annealing reaction was analyzed: 3uM of NC was added to TAR and cTAR (each 1uM) prefolded in TNMg 1X at r.t. for different incubation times (2, 15, 30, 45, 60, 90, 120 minutes). We can observe in Figure 1.35b that after 45 minutes at r.t. the annealing reaction (oligos:NC=1:3) is completed.

In view of the application of this electrophoresis assay to investigate the effects of compounds on the TAR/cTAR annealing reaction, some adjustments to the protocol was introduced. When introduce compounds in the gel system, a short time of incubation with the protein is necessary to prevent nucleic acids aggregation and precipitation. For this reason we

used the molar ratio oligos:NC of 1:8 to pick out the shortest time necessary to form the hybrid. The samples were resolved by 12% non-denaturing polyacrylamide gels containing 1X TBE buffer, to allow the protein to enter into the gel. The result of this experiment was reported in Figure 1.36.

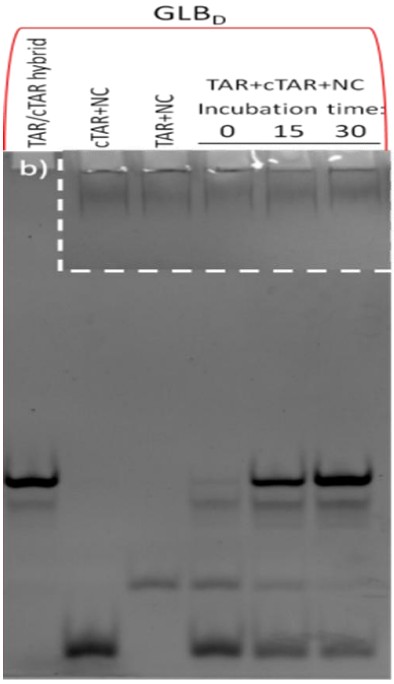


Figure 1.36 Improvement of the analysis of the TAR/cTAR annealing reaction in the presence of the full-length NC protein. Controls: hybrid TAR/cTAR, TAR+NC and cTAR+NC (each 1uM) folded in TNMg 1X. TAR and cTAR, prefolded in TNMg 1X, were incubated with NC 8µM (oligos:NC=1:8) for 0, 15' or 30' at room temperature. The samples were added of denaturing Gel Loading Buffer. Electrophoresis on a 20% polyacrylamide gel at r.t. in TBE 1X. After electrophoresis nucleic acids on the gel were stained with EtBr; nucleic acids fluorescence in the gel were detected on a Geliance 600 Imaging System **(a)**; NC protein was stained with Coomassie Blue dye **(insert b)**.

We can observe that the band of the hybrid formed by NC had the same mobility of the heat-annealed hybrid and therefore the annealing of TAR with its complementary sequence cTAR induced by NC is verified. Using the oligos:NC ratio 1:8, at t=0 minutes incubation time the annealing reaction was already started. The hybrid formation was completed at 30 minutes since the intensity of the cTAR band was almost disappeared while at 15 minutes the reaction is already progressing. We could notice also that as the time went by the intensity of the hybrid band increased while the intensity of the single stranded nucleic acids bands reduced. Finally the protein in the gel was stained with Coomassie dye to verify that NC entered into the gel (Figure 1.36 insert b). Hence, we can identify the optimal conditions to form the TAR/cTAR hybrid in a short time: after 15 minutes at r.t. the annealing reaction (oligos:NC=1:8) is formed.

1.4.2 Identification of potential NC inhibitors

1.4.2.1 High Throughput Screening (HTS) of NC inhibitors

After the thorough analysis of the full-length protein NC and its interaction with viral nucleic acids, the effect of compounds on the chaperone activity of NC was investigated. A simple, fast and highly reproducible FRET assay for the High Throughput Screening of inhibitors of the NC nucleic acids destabilization activity, in order to identify potential antiviral agents, was performed and optimized.

HTS allows, as already reported in the literature [61], the analysis of libraries of small molecules able to inhibit NC chaperone activity and to determine the IC_{50}. More of 100 molecules from our in-house chemical library, representing a wide range of structures with potential pharmacological activity, were analyzed by HTS. After the screening of all compounds, inhibition values were used to classify all the compounds in the following group: non active (IC_{50} >50μM), barely active (50μM≥ IC_{50} >20μM), moderately active (20μM≥ IC_{50} >5μM) and highly active (IC_{50} ≤5μM).

Impairment of NC chaperone activity on cTAR and TAR i.e. of the cTAR and TAR melting was monitored by adding increasing concentrations of compounds: in this experiment the compounds ability to reduce the NC chaperone activity was evaluated according to oligonucleotides fluorescence emission (λ_{exc}=485nm and λ_{em}=535nm). If NC activity is reduced, the proximity of cTAR (or TAR) ends induces a strong fluorescence quenching. NC is able to melt the lower half of the oligonucleotide stem increasing the distance between the two dyes and eliciting a raise of FAM fluorescence.

In this work we focus on the investigation of the 2,6-dipeptidyl anthraquinones belonging to Z, GSF and G series. The results of the HTS are summarized in Table 1.1, in which are reported the IC_{50} value of each tested compound. The half maximal inhibitory concentration (IC_{50}) is the measure of the effectiveness of a compound in inhibiting the chaperone activity of NC on TAR and on cTAR.

We verified compounds fluorescence exciting at λ=485nm and recording fluorescence emission at λ=535 nm. None of the tested compounds resulted to emit at λ=485nm and none was found to affect the fluorescence of the doubly-labeled nucleic acids (data not shown). These results confirmed that background fluorescence in the plate reader was not due to compounds, hence, it permitted to avoid false negatives and also false positives, assuring that the reduction of fluorescence was not due to a quenching process between FAM and compound but to the inhibition of NC.

compound	IC_{50} (uM) NC-TAR	IC_{50} (uM) NC-cTAR	compound	IC_{50} (uM) NC-TAR	IC_{50} (uM) NC-cTAR
Z6	4,49±0,31	2,07±0,02	Z9	2,60±0,04	2,28±0,22
Z7	4,17±0,13	2,22±0,05	Z8	5,85±0,11	3,87±0,41
542	3,25±0,13	6,42±0,25	721	3,71±0,15	3,84±0,14
543	4,43±0,44	4,66±0,28	723	12,00±0,65	10,28±0,12

compound	IC_{50} (uM) NC-TAR	IC_{50} (uM) NC-cTAR	compound	IC_{50} (uM) NC-TAR	IC_{50} (uM) NC-cTAR
GSF1	10,04±0,45	11,8±0,84	GSF8	21,39±1,36	22,76±3,08
GSF3	17,63±0,57	13,66±1,2	GSF12	13,79±0,69	15,14±1,57
GSF6	3,2±0,1	2,04±0,03	GSF13	42,88±3,7	36,27±2,84
GSF7	6,3±0,23	7,1±0,89			

compound	IC_{50} (uM) NC-TAR	IC_{50} (uM) NC-cTAR	compound	IC_{50} (uM) NC-TAR	IC_{50} (uM) NC-cTAR
G1	12,37±0,53	11,96±0,13	G12	11,74±1,31	4,37±0,38
G3	27,40±0,72	19,65±1,11	G13	18,04±0,99	12,28±0,61

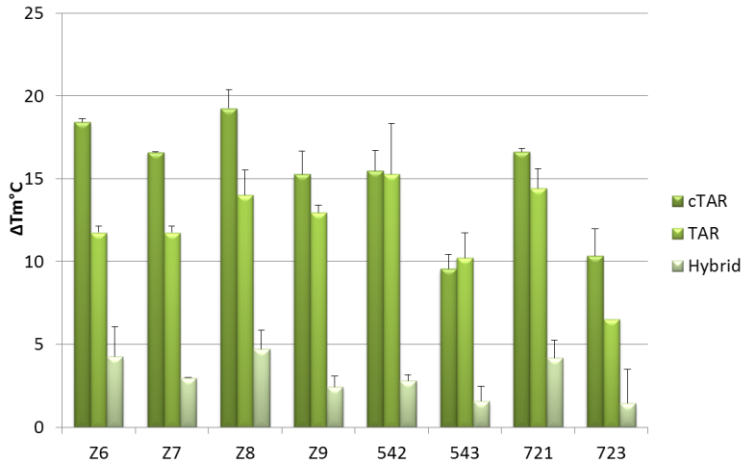
Table 1.1 IC_{50} value of each compound in inhibiting the chaperone activity of NC on TAR and on cTAR obtained by HTS.

We can observe that all tested anthraquinones are active both toward TAR and cTAR. Compounds of series Z can stabilize with high activity both TAR and cTAR structures but they have a slight preference for cTAR. Only compound 542 can stabilize mostly the stem-loop structure of TAR compared to cTAR. All anthraquinones Z result high active potential inhibitors of NC biological activity. Compounds of series GSF and G can be divided in highly-, moderately- and barely-active NC inhibitors. We can identify the most promising compounds that are GSF6 and GSF7; compounds G12 is attractive to study because it shows a high activity as inhibitor of NC toward cTAR but is a medium-active inhibitor toward TAR.

1.4.2.2 Evaluation of the mechanism of action of 2,6-dipeptidyl anthraquinones: analysis of oligonucleotides-compounds interactions

Compounds could inhibit NC activity through different mechanisms of action. Anthraquinones are known DNA and RNA binders and therefore it is reasonable to associate their NC chaperone activity impairment with the stabilization of nucleic acid stem-loop structure. Melting Assay allowed us to estimate the compounds capability to alter nucleic acids thermal denaturing profile, due to direct interaction and structure stabilization. This assay allowed the understanding of the mechanism of action of NC-inhibiting compounds. If we calculate a positive ΔTm value it means that the compound interacts with nucleic acid, probably by intercalation, restraining melting. Melting Assay was performed as described in the Experimental Section. First, the fluorescence emission of the oligonucleotides in the presence of the maximum concentration of DMSO tested for each compound (1% v/v) was monitored, from 25 to 99°C, and we could verify that DMSO did not cause any significant fluorescence reduction. The reference Tm values, in the absence of the protein, resulted to be: 53,80 °C for cTAR; 69,30 °C for TAR and 69,40 °C for the DNA/RNA hybrid. ΔTm values at 10μM compound concentration resulted to be the most significant and are reported below (Figure 1.37, 1.38 and 1.39).

Z compounds



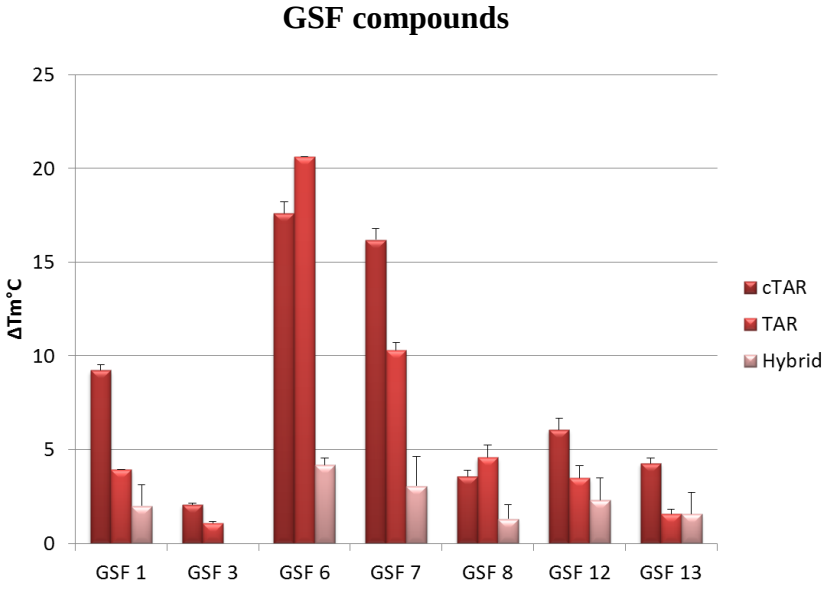
Comp	ΔTm°C cTAR	ΔTm°C TAR	ΔTm°C Hybrid	Comp	ΔTm°C cTAR	ΔTm°C TAR	ΔTm°C Hybrid
Z6	18,41±0,21	11,75±0,39	4,25±1,8	**542**	15,49±1,2	15,29±3,04	2,81±0,36
Z7	16,59±0,04	11,75±0,38	2,98±0,01	**543**	9,56±0,86	10,24±1,49	1,57±0,91
Z8	19,26±1,11	14,03±1,49	4,72±1,14	**721**	16,64±0,18	14,41±1,20	4,17±1,07
Z9	15,27±1,41	12,96±0,43	2,43±0,65	**723**	10,33±1,64	6,53±0,05	1,44±2,04

Figure 1.37 a. Graphic representation of ΔTm values of cTAR, TAR and TAR/cTAR hybrid in TNMg 1X at 10μM anthraquinones concentration. **b.** ΔTm values of cTAR, TAR and TAR/cTAR hybrid in TNMg 1X at 10μM anthraquinones concentration.

At 1μM compound concentration the stabilization was not so evident and did not allow to analyze properly the stabilizing power of the compounds while at 100μM ΔTm was either too high or compounds tended to precipitate.

All tested 2,6-dipeptidyl anthraquinones can stabilize dose-dependently the bulge-loop structure of both TAR and cTAR. The Figure 1.37 reports the results obtained at AQ/nucleic acid=10. ΔTm of the TAR/cTAR duplex is low, meaning that the interaction with the hybrid is less significant. Anthraquinones derivatives recognize selectively TAR instead of not specific DNA [62].

These data supported HTS results since all of the compounds of Z series proved to be highly active NC inhibitors. In the case of Z9, for example, had an IC_{50} value of 2.6μM on TAR and 2.28μM on cTAR. This was a concrete evidence that the strategy of inhibit NC protein through stem-loop stabilization can be followed.

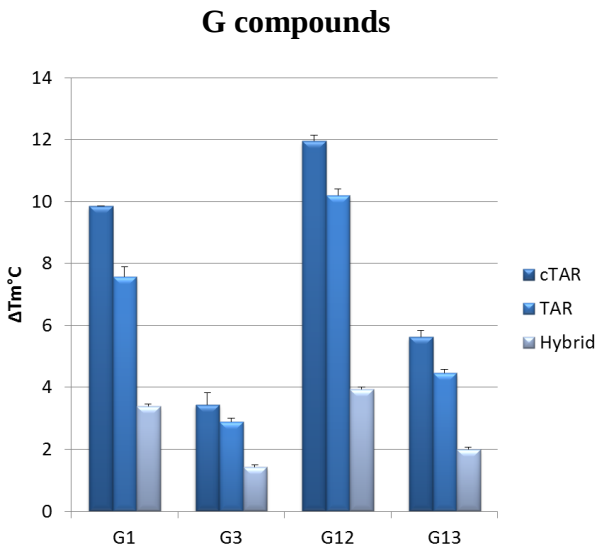


Comp	ΔTm°C cTAR	ΔTm°C TAR	ΔTm°C Hybrid	Comp	ΔTm°C cTAR	ΔTm°C TAR	ΔTm°C Hybrid
GSF1	9,23±0,3	3,95±0,00	2,00±1,13	**GSF8**	3,57±0,32	4,59±0,66	1,31±0,77
GSF3	2,07±0,06	1,07±0,08	0	**GSF12**	6,07±0,61	3,48±0,67	2,30±1,19
GSF6	17,62±0,59	20,64 (*)	4,17±0,36	**GSF13**	4,27±0,27	1,58±0,24	1,56±1,15
GSF7	16,20±0,59	10,32±0,37	3,06±1,56				

Figure 1.38 a. Graphic representation of ΔTm values of cTAR, TAR and TAR/cTAR hybrid in TNMg 1X at 10μM anthraquinones concentration. **b.** ΔTm values of cTAR, TAR and TAR/cTAR hybrid in TNMg 1X at 10μM anthraquinones concentration. (*) measure obtained at 5 μM GSF6.

Figure 1.38 reports the results obtained at GSF/nucleic acid=10. At 10 and 100μM GSF6 concentrations we were not able to measure the melting temperature of TAR while at 1μM it was possible. Therefore we reduced the concentration of the tested compound from 10 to 5μM

and measured the ΔTm. Only 2,6-peptidyl anthraquinones GSF6 and GSF7 can significantly stabilize dose-dependently the bulge-loop structure of both TAR and cTAR, as observed for positive control Z9. The other derivatives cannot stabilize these nucleic acid structures, except GSF1 which shows a significant ΔTm with cTAR. ΔTm of the TAR/cTAR duplex is lower, meaning that the interaction with the hybrid is less significant. GSF6 and GSF7 result to be the most able to stabilize viral nucleic acids and were also highly active inhibitors of NC chaperone activity. GSF6 for instance had IC_{50} values of 2.04 on cTAR and of 3.02 on TAR. Also for GSF anthraquinones we could affirm that melting assay data supported HTS results.



Comp	ΔTm°C cTAR	ΔTm°C TAR	ΔTm°C Hybrid
G1	9,85 ± 0,01	7,58 ± 0,32	3,39 ± 0,06
G3	3,44 ± 0,39	2,89 ± 0,11	1,43 ± 0,06
G12	11,95 ± 0,20	10,19 ± 0,22	3,95 ± 0,06
G13	5,63 ± 0,20	4,46 ± 0,13	2,00 ± 0,07

Figure 1.39 a. Graphic representation of ΔTm values of cTAR, TAR and TAR/cTAR hybrid in TNMg 1X at 10μM anthraquinones concentration. **b.** ΔTm values of cTAR, TAR and TAR/cTAR hybrid in TNMg 1X at 10μM anthraquinones concentration.

Figure 1.39 reports the results obtained at G/nucleic acid=10. The G series anthraquinones can interfere more with cTAR compared to TAR. In particular, G1 and G12 show a significant ΔTm with TAR and cTAR. ΔTm of the TAR/cTAR duplex are lower, meaning that the interaction with the hybrid is less important. G12 compound results to be the most able to stabilize viral nucleic acids and from HTS it resulted to be highly active as inhibitor of NC chaperone activity, having IC_{50} values of 4.37 on cTAR and of 11.74 on TAR. Melting assay data were in agreement with HTS results also for anthraquinones of G series.

1.4.2.3 Analysis of anthraquinones effect on the TAR/cTAR annealing reaction

We have previously showed that NC dramatically catalyzes in vitro the annealing of cTAR with TAR, and the reaction can be monitored by gel electrophoresis. The conditions suitable for test compounds were identified (see paragraph 1.4.1.3); therefore we decided to evaluate the effect of anthraquinones on the NC-induced hybrid formation using the same gel system: increasing the concentrations (2, 20, 100 and 200μM) of the most active anthraquinones resulting from HTS we evaluated the inhibition of NC annealing activity through native gel electrophoresis. The aim of the experiment was also to understand if there is a correlation between the inhibition of TAR/cTAR melting and the inhibition of the annealing.

For these experiments unlabeled TAR and cTAR were used. TAR and cTAR were incubated separately with increasing concentration of compound for 15 minutes, were then mixed and incubated with the protein for other 15 minutes. This time was enough for NC to form the hybrid. Another particular step was introduced in the protocol to test compounds: the samples were kept steady on ice until they were loaded into the gel. This step was introduced to prevent the nonspecific annealing of TAR and cTAR (already melted by NC, but not annealed yet) through Kissing pathway and to quench the hybrid formation reaction just when NC is removed from the complex with the oligonucleotides. The samples were analyzed by gel electrophoresis. The images of the assays are reported below. In figure 1.40 are shown the results of anthraquinones Z9 (a) and Z8 (b).

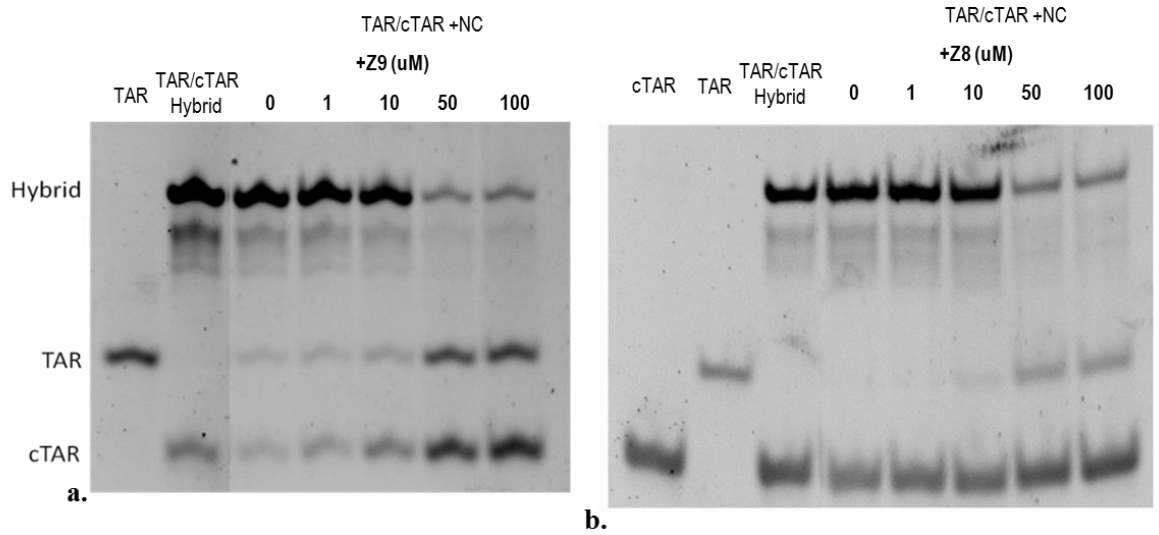


Figure 1.40 Inhibition of the TAR/cTAR annealing reaction by Z9 (a) and Z8 (b) anthraquinones. Controls: TAR, cTAR and the hybrid TAR/cTAR (each 1uM) folded in TNMg 1X; TAR and cTAR were incubated separately with increasing concentration of compound (0, 1, 10, 50, 100uM) for 15 minutes at room temperature. TAR and cTAR pre-incubated with the compound were then mixed and incubated with NC 8uM (oligos:NC=1:8) for 15 minutes at room temperature. The samples were added of denaturing Gel Loading Buffer. The samples were kept steady on ice until they were loaded into the gel. The samples were analyzed by

electrophoresis on a 12% polyacrylamide gel at r.t. in TBE 1X. After electrophoresis nucleic acids on the gel were stained with SybrGreen II.

Figure 1.40a shows the analysis of Z9, resulted to be highly active both on TAR and on cTAR by HTS, and Figure 1.40b shows the results of Z8, moderately active on TAR and highly active on cTAR by HTS. Z9 compound results highly active in reducing the hybrid formation: at 50 and 100μM compound concentrations the hybrid band is almost disappeared. On the contrary, the intensity of cTAR and TAR bands increases proportionally with compound concentration. Anthraquinone Z8, like Z9, shows a good inhibitory activity on hybrid formation.

In Figure 1.41a is reported the result of anthraquinone 542 (highly active both on TAR and cTAR by HTS), in Figure 1.41b of 543 (highly active on TAR and moderately on cTAR by HTS) and of 721 (highly active on TAR and cTAR by HTS) (Figure 1.41c).

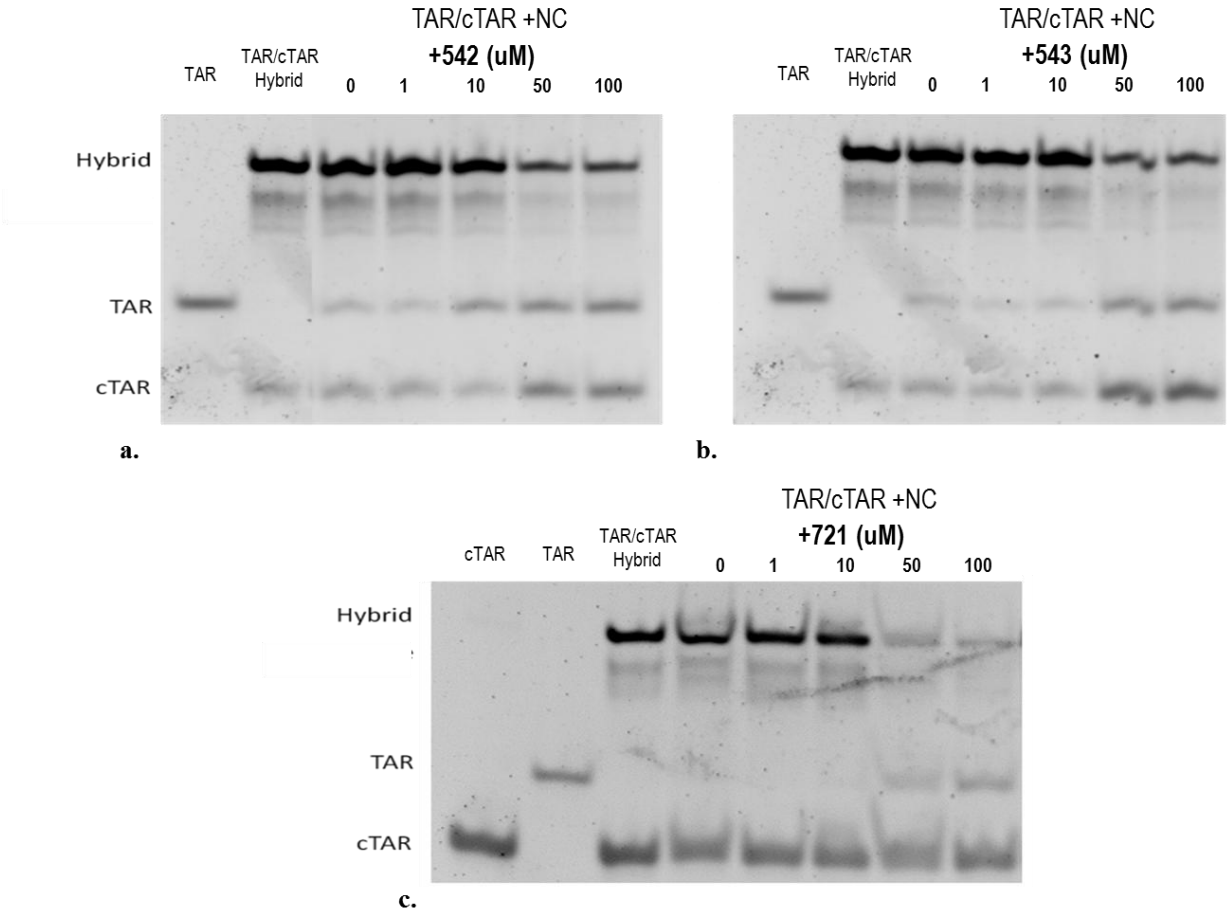


Figure 1.41 Inhibition of the TAR/cTAR annealing reaction by 542 (a), 543 (b) and 721 (c) anthraquinones. Controls: TAR, cTAR and the hybrid TAR/cTAR (each 1uM) folded in TNMg 1X; TAR and cTAR were incubated separately with increasing concentration of compound (0, 1, 10, 50, 100uM) for 15 minutes at room temperature. TAR and cTAR pre-incubated with the compound were then mixed and incubated with NC 8uM (oligos:NC=1:8) for 15 minutes at room temperature. The samples were added of denaturing Gel Loading Buffer. The samples were kept steady on ice until they were loaded into the gel. The samples were

analyzed by electrophoresis on a 12% polyacrylamide gel at r.t. in TBE 1X. After electrophoresis nucleic acids on the gel were stained with SybrGreen II.

721, 542 and 543 compounds result to be highly active in reducing the hybrid formation, at 50 and 100μM compounds concentrations the intensity of the hybrid bands are almost disappeared. On the contrary, the intensity of TAR and cTAR bands increases proportionally with compound concentration.

GSF6 (highly active on TAR and cTAR by HTS), GSF7 (moderately active on TAR and cTAR by HTS) and GSF13 (barely active on TAR and cTAR by HTS) were tested and results are shown in Figure 1.42a, b, and c respectively.

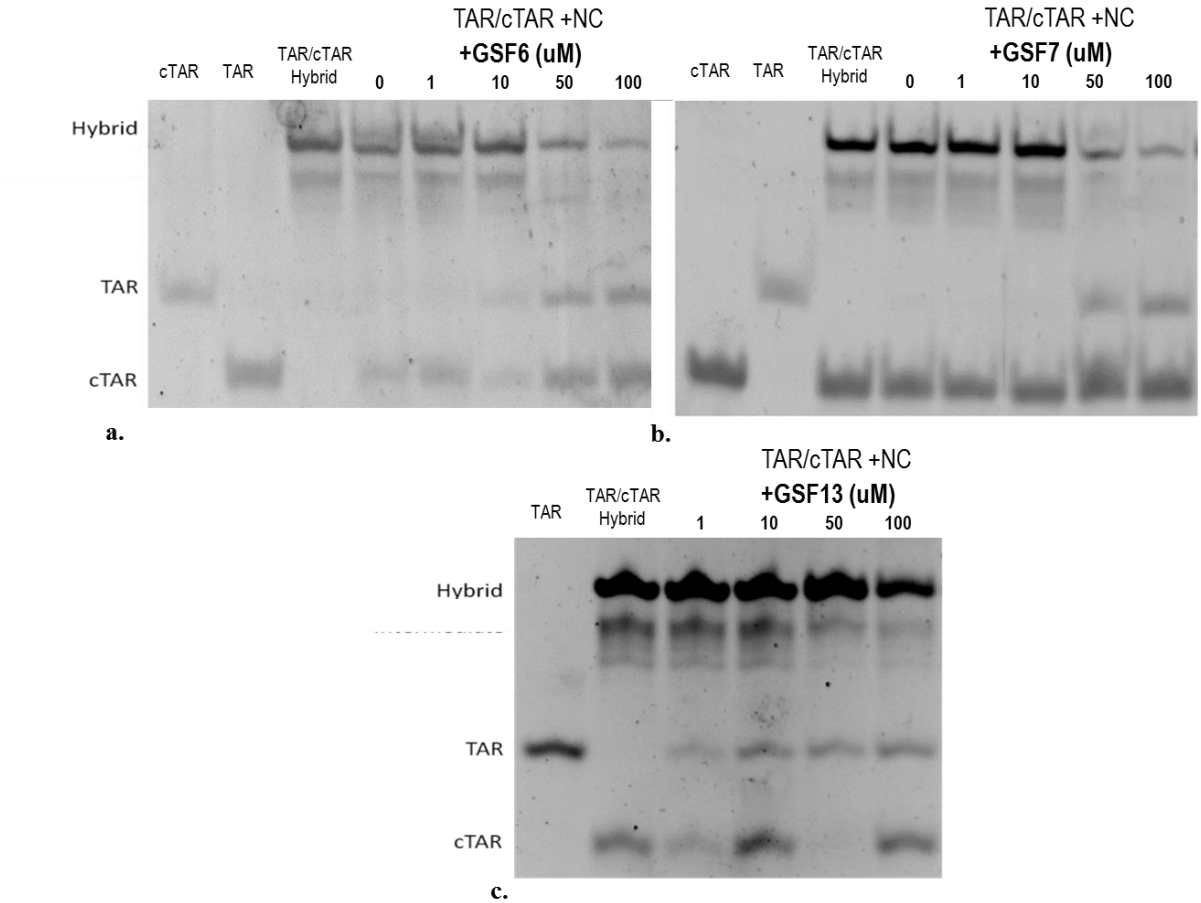


Figure 1.42 Inhibition of the TAR/cTAR annealing reaction by GSF6 (a), GSF7 (b) and GSF13 (c) anthraquinones. Controls: TAR, cTAR and the hybrid TAR/cTAR (each 1uM) folded in TNMg 1X; TAR and cTAR were incubated separately with increasing concentration of compound (1, 10, 50, 100uM) for 15 minutes at room temperature. TAR and cTAR pre-incubated with the compound were then mixed and incubated with NC 8uM (oligos:NC=1:8) for 15 minutes at room temperature. The samples were added of denaturing Gel Loading Buffer. The samples were kept steady on ice until they were loaded into the gel. The samples were analyzed by electrophoresis on a 12% polyacrylamide gel at room temperature in TBE 1X. After electrophoresis nucleic acids on the gel were stained with SybrGreen II.

Both GSF7 and GSF6 result to be highly active and are able to completely impair the hybrid formation at 50 and 100μM concentrations. The intensity of TAR and cTAR bands instead increases proportionally with compound concentration as expected. GSF13 reduces a little bit the hybrid formation at 100μM compound concentration.

Finally, G12 which resulted to be moderately active on TAR and highly active on cTAR by HTS was tested (Figure 1.43).

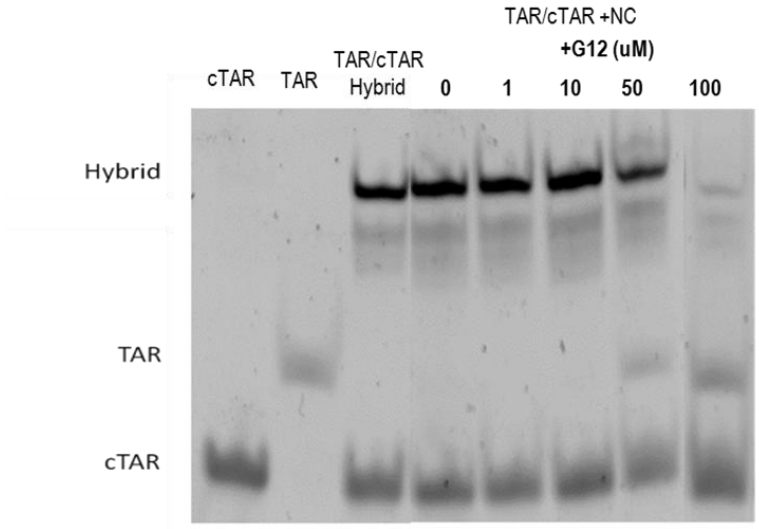


Figure 1.43 Inhibition of the TAR/cTAR annealing reaction by GSF12 anthraquinone. Controls: TAR, cTAR and the hybrid TAR/cTAR (each 1uM) folded in TNMg 1X; TAR and cTAR were incubated separately with increasing concentration of compound (0, 1, 10, 50, 100uM) for 15 minutes at room temperature. TAR and cTAR pre-incubated with the compound were then mixed and incubated with NC 8uM (oligos:NC=1:8) for 15 minutes at room temperature. The samples were added of denaturing Gel Loading Buffer. The samples were kept steady on ice until they were loaded into the gel. The samples were analyzed by electrophoresis on a 12% polyacrylamide gel at room temperature in TBE 1X. After electrophoresis nucleic acids on the gel were stained with SybrGreen II.

G12 results to be able to reduce the hybrid formation at 100μM concentration. The intensity of cTAR band increases proportionally with compound concentration as expected.

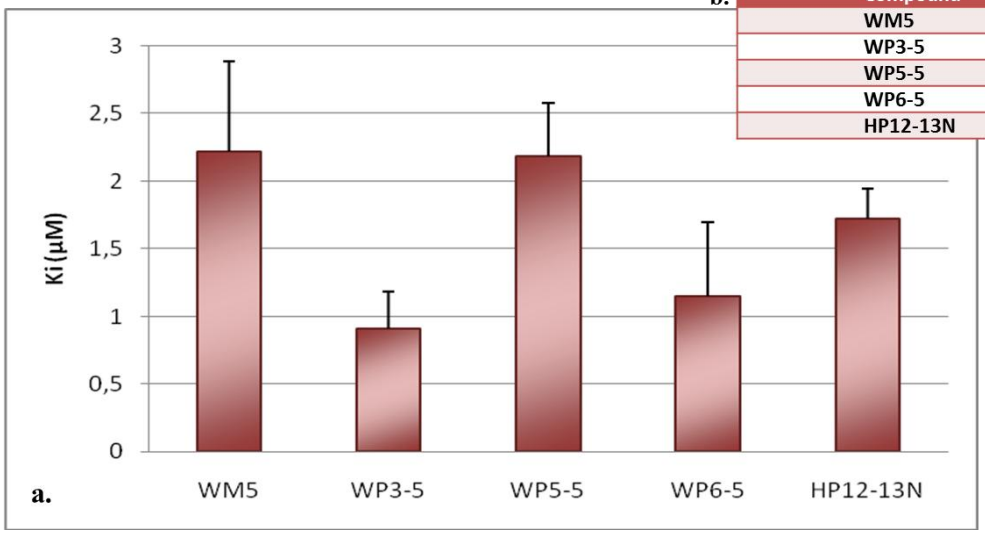
All the tested anthraquinones result to be inhibitors of the TAR/cTAR hybrid formation, according with the NC inhibitory activity analyzed by HTS. Hence, there is a correlation between the inhibition of TAR/cTAR melting and the inhibition of the annealing.

1.4.3 Identification of inhibitors of the Tat-mediated transcription process

1.4.3.1 Analysis by FQA of the quinolones interference on the Tat-TAR complex

In the past few years, a fast and versatile platform of screening to examine the interference of compounds against the Tat-TAR complex was performed and validated [50]. Fluorescence Quenching Assay (FQA) allows, as already reported in the literature [49], a fast analysis of libraries of small molecules able to compete with the Tat-TAR complex and to determine the inhibition constant Ki. FQA is based on FRET (Förster Resonance Energy Transfer) process in which the donor molecule is represented by the fluorophore fluorescein linked at the N-terminus of the 10aa Tat peptide, while the acceptor group is the quencher dabcyl linked to the 3'-end of TAR-RNA. The energy transfer occurs after the Tat-TAR complex formation and its disruption by compounds leads to a decrease of the energy transfer, resulting in the increase of fluorescence intensity. Titrations of labeled Tat/TAR in the presence of competitors were performed, as reported in the Experimental Section, in the plate reader Victor III; the fluorescein-labeled peptide was excited at 490 nm and emission was recorded at 535 nm. Fluorescence intensities in the presence of compounds were measured and curves were fitted to obtain Ki values.

The Ki values of tested quinolones are reported in Figure 1.44.



b.

Compound	Ki Tat/TAR (µM)
WM5	2.22 ± 0.66
WP3-5	0.91 ± 0.27
WP5-5	2.18 ± 0.40
WP6-5	1.15 ± 0.55
HP12-13N	1.72 ± 0.22

Figure 1.44 (a) Graphic representation of Ki values of tested quinolones. (b) Ki values of tested 6-aminoquinolones. Values obtained by FQA: TAR dabcyl was titrated in the presence of 10 nM fluorescein-Tat peptide (10aa) and 10µM compound in TNMg 1X. WM5 was used as control.

Both the WP and HP quinolones series resulted to be unstable in solution over time. To perform the FRET assay, the compound has to be stable in the aqueous buffer for almost 10-15 minutes: after having characterized their stability in aqueous buffer by spectrophotometric analysis, WP3-5, WP5-5, WP6-5 and HP12-13N were tested by FQA, using WM5 as positive control [49].

We can observe that all the tested 6-DFQs of WP series are active as Tat-TAR complex formation inhibitors, with Ki values lower than the control WM5.

The 6-DFQs of the WP series differ from the control for modifications introduced at the N1 position: these modifications do not result in loss of activity, but rather seem to improve the activity of WP quinolones as inhibitors of the Tat-TAR complex formation.

The only compound of the HP series that has been possible to study is HP12-13N, which results to be active as inhibitors of the Tat-TAR complexation.

1.4.3.2 Analysis by EMSA of the ELs interference on the Tat-TAR complex

ELs were tested as inhibitors of the Tat-TAR complex formation. ELs analysis was not possible to be performed by FQA due to a quenching effect observing after the addition of ELs to fluoresceinated Tat. The analysis of the ELs interference on the Tat-TAR complex was therefore performed by EMSA.

The Electrophoresis Mobility Shift Assay (EMSA) allows the evaluation of the in vitro activity of compounds to interfere with the Tat-TAR complex formation. This assay is based on the different electrophoretic mobility of the free and complexed TAR, and in particular an experimental protocol that uses the staining with the fluorescent molecule SybrGreen II was performed and validated in our previous research [50].

With this method the band of the Tat-TAR complex is not detectable, because, when TAR is bound to Tat, the dye cannot intercalate RNA; however, the complexation reaction can be monitored observing the disappearance of the band corresponding to free TAR. The addition of increasing concentrations of compound determines, if the compound impairs the complex formation, the re-appearance of the free TAR band.

Tat 10aa was incubated with increasing concentrations of each compound at room temperature for 15 minutes, then folded TAR was added for other 15’ minutes. The gel shifts obtained are reported below (Figure 1.45, 1.46, 1.47, 1.48, 1.49 and 1.50).

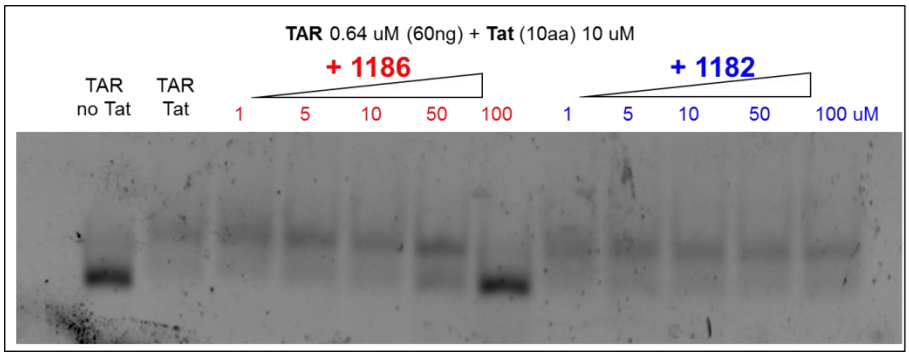


Figure 1.45 EMSA to evaluate the inhibition of the Tat-TAR complex formation by 1186 (on the left) and 1182 (on the right) compounds. Controls: 60ng TAR folded in TNMg 1X in the presence and in the absence of Tat 10aa (10µM). the Tat-TAR complex formation was analyzed in the presence of increasing concentrations (1, 5, 10, 50, 100 µM) of compound. Electrophoresis on a 10% polyacrylamide gel in TBE 1X. After electrophoresis nucleic acids on the gel were stained with SybrGreenII; nucleic acids fluorescence in the gel were detected on a Geliance 600 Imaging System (PerkinElmer).

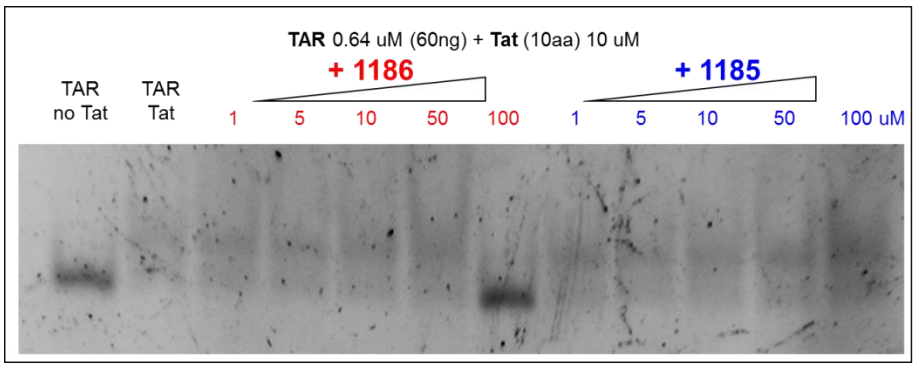


Figure 1.46 EMSA to evaluate the inhibition of the Tat-TAR complex formation by 1186 (on the left) and 1185 (on the right) compounds. Controls: 60ng TAR folded in TNMg 1X in the presence and in the absence of Tat 10aa (10µM). the Tat-TAR complex formation was analyzed in the presence of increasing concentrations (1, 5, 10, 50, 100 µM) of compound. Electrophoresis on a 10% polyacrylamide gel in TBE 1X. After electrophoresis nucleic acids on the gel were stained with SybrGreenII; nucleic acids fluorescence in the gel were detected on a Geliance 600 Imaging System (PerkinElmer).

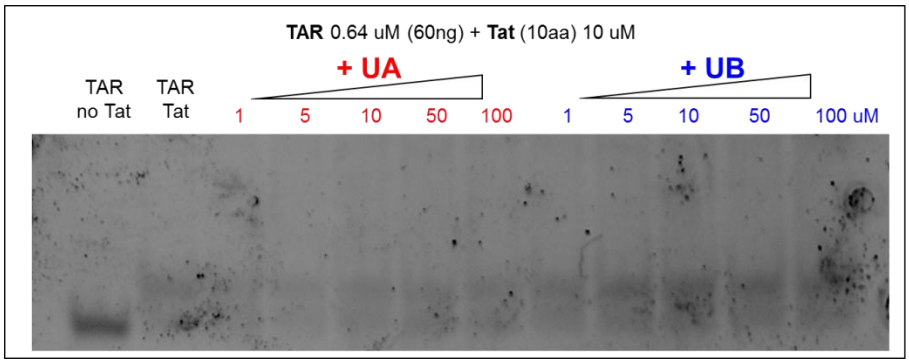


Figure 1.47 EMSA to evaluate the inhibition of the Tat-TAR complex formation by UA (on the left) and UB (on the right) compounds. Controls: 60ng TAR folded in TNMg 1X in the presence and in the absence of Tat 10aa (10µM). the Tat-TAR complex formation was analyzed in the presence of increasing concentrations (1, 5, 10, 50, 100 µM) of compound. Electrophoresis on a 10% polyacrylamide gel in TBE 1X. After electrophoresis nucleic acids on the gel were stained with SybrGreenII; nucleic acids fluorescence in the gel were detected on a Geliance 600 Imaging System (PerkinElmer).

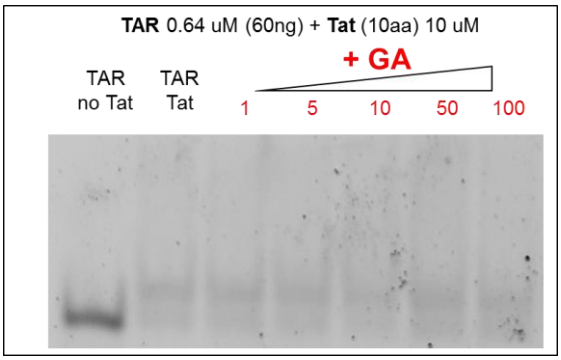


Figure 1.48 EMSA to evaluate the inhibition of the Tat-TAR complex formation by GA. Controls: 60ng TAR folded in TNMg 1X in the presence and in the absence of Tat 10aa (10µM). the Tat-TAR complex formation was analyzed in the presence of increasing concentrations (1, 5, 10, 50, 100 µM) of compound. Electrophoresis on a 10% polyacrylamide gel in TBE 1X. After electrophoresis nucleic acids on the gel were stained with SybrGreenII; nucleic acids fluorescence in the gel were detected on a Geliance 600 Imaging System (PerkinElmer).

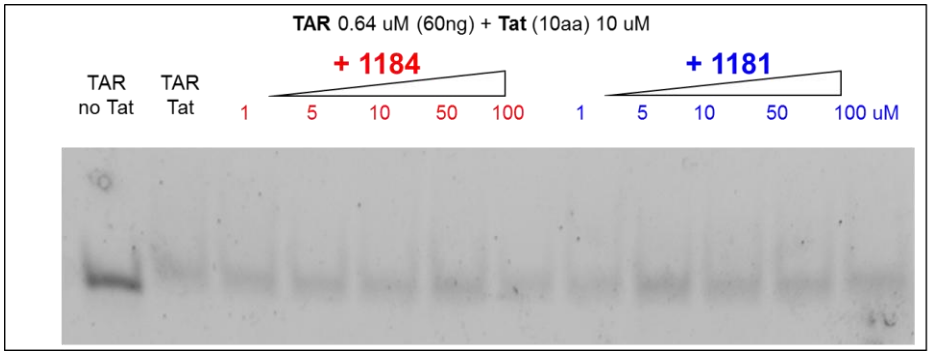


Figure 1.49 EMSA to evaluate the inhibition of the Tat-TAR complex formation by 1184 (on the left) and 1181 (on the right) compounds. Controls: 60ng TAR folded in TNMg 1X in the presence and in the absence of Tat 10aa (10µM). the Tat-TAR complex formation was analyzed in the presence of increasing concentrations (1, 5, 10, 50, 100 µM) of compound. Electrophoresis on a 10% polyacrylamide gel in TBE 1X. After electrophoresis nucleic acids on the gel were stained with SybrGreenII; nucleic acids fluorescence in the gel were detected on a Geliance 600 Imaging System (PerkinElmer).

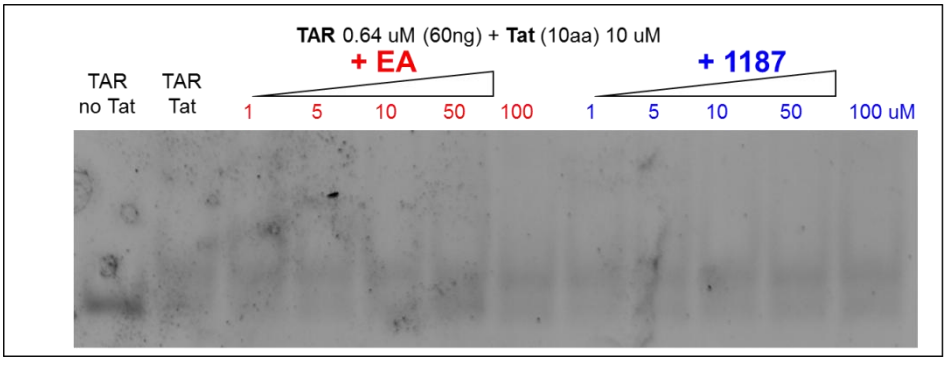


Figure 1.50 EMSA to evaluate the inhibition of the Tat-TAR complex formation by EA (on the left) and 1187 (on the right) compounds. Controls: 60ng TAR folded in TNMg 1X in the presence and in the absence of Tat 10aa (10µM). the Tat-TAR complex formation was analyzed in the presence of increasing concentrations (1, 5, 10, 50, 100 µM) of compound. Electrophoresis on a 10% polyacrylamide gel in TBE 1X. After electrophoresis nucleic acids on the gel were stained with SybrGreenII; nucleic acids fluorescence in the gel were detected on a Geliance 600 Imaging System (PerkinElmer).

In all the experiments, the Tat-TAR complex formation is verified: in the presence of Tat, the free TAR band disappears and is replaced by a weakly detectable smeared band.
1182 compound (Figure 1.45, on the right) and 1185 compound (Figure 1.46, on the right) are inactive for the inhibition of the Tat-TAR complex formation.

In Figure 1.47 are reported the results of UA (on the left) and of UB (on the right): they are inactive.

GA (Figure 1.48) and EA (Figure 1.50, on the left) do not improve the Tat-TAR complex formation.

In Figure 1.49 are reported the analysis of 1184 (on the left) and of 1181 (on the right) that result to have no activity as Tat-TAR complexation inhibitors.

In Figure 1.45 (on the left) and in Figure 1.46 (on the left) are reported the gel shifts in which 1186 compound was analyzed. The addition of 100 μM of 1186 compound determines the completely re-appearance of the free TAR band. In the presence of high concentration of 1186 the Tat-TAR complex does not occur. Only compound 1186 can act as Tat-TAR complex formation inhibitor.

1.5 CONCLUSIONS

NC recognition and binding of viral nucleic acids represents a new molecular approach for antiretroviral therapy [63, 64], and TAR and cTAR are a valid target for the development of new compounds whose interactions with nucleic acids would lead to the inhibition of the NC-mediated functions. Small molecules that are able to recognize these viral nucleic acids were considered in this thesis as potential inhibitors of reverse transcription of HIV-1 for their inhibition of the NC protein activity.

The HIV-1 full-length protein was deeply studied prior to investigate the effect of compounds on its chaperone activity. The characterization of NC structure was completed by circular dichroic spectroscopy both in the presence and absence of zinc ions, confirming its Zn-finger structure.

The interaction of both TAR and cTAR viral nucleic acids with the full-length NC was investigated by a simple fluorescence assay. We can affirm that the biological activity of the full-length NC is verified and the conditions in which the protein is active as nucleic acids chaperone on both TAR and cTAR are identified. By the fluorescence assay, we have also confirmed that the nucleic acids destabilization activity of NC resides in the Zn finger structures and the Zn finger structures result therefore necessary and indispensable for the chaperone activity of NC, since the unfolded NC was not active. Moreover, the fluorescence assay for the analysis of the biological activity of NC allowed us to understand that Zn ions in solution determine a quenching effect on FAM fluorescence, according with literature data [58]. NC biological activity was analyzed on TAR and cTAR making also use of a microplate reader, which allows the use of lower volumes of sample. In all cases it was showed that the stem-structure of TAR results more stable of the cTAR stem, according to the results obtained by the fluorescence assay. NC results therefore active as chaperone on both TAR and cTAR, but results more active on cTAR. NC can destabilize more efficiently DNA compared with RNA sequences.

Furthermore, under conditions previously identified by fluorescence assays, the annealing of TAR and cTAR in the presence of NC was analyzed by gel-electrophoresis. In the absence of NC the annealing reaction is extremely slow, while the hybrid TAR/cTAR formation is dramatically accelerated by NC. We have identified the lower NC concentration necessary to obtain the complete hybrid TAR/cTAR formation and the kinetics of the annealing reaction was analyzed. In view of the application of this electrophoresis assay to investigate the effects of compounds on the TAR/cTAR annealing reaction, the optimal conditions to test compounds were identified.

After the thorough analysis of the full-length NC protein and its interaction with viral nucleic acids, the effect of compounds on the NC chaperone activity was investigated. We have devised a FRET-based strategy to screen our wide in-home chemical library of small molecules: a simple, fast and highly reproducible High Throughput Screening (HTS) of potential NC inhibitors was optimized. Besides what specifically shown in this thesis, more than 100 molecules, representing a wide range of structures, were tested by HTS in our lab with my collaboration and under my supervision [65]. The general results of the HTS, summarized in Figure 1.51, allowed the division of the screened compounds into four main groups by their ability to inhibit the NC chaperone activity on TAR and on cTAR (based on IC_{50} values).

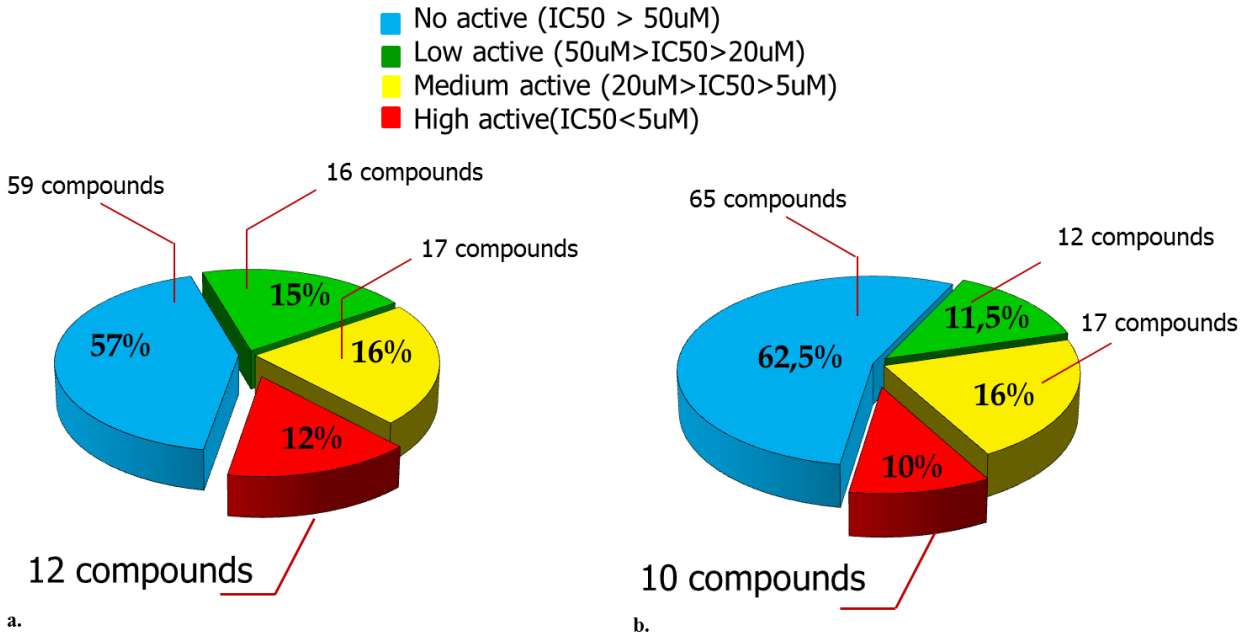


Figure 1.51 Results of HTS: IC_{50} distribution of tested compounds on cTAR (a) and on TAR (b).

We founded about 10% of highly active and 16% of medium active molecules. All 2,6-dipetydil anthraquinones of series Z, GSF and G analyzed in this work resulted to be active as NC inhibitors both on TAR and cTAR, and they can be classified as reported in Table 1.2.

		TAR	cTAR
Low activity	**50≥IC_{50}>20uM**	GSF8, GSF13, G3	GSF8, GSF13
Medium activity	**20≥IC_{50}>5uM**	Z8, 723, GSF1, GSF3, GSF7, GSF12, G1, G12, G13	542, 723, GSF1, GSF3, GSF7, GSF12, G1,G3, G13
High activity	**IC_{50}≤5uM**	Z6, Z7, Z9, 542,543,721; GSF6	Z6, Z7, Z8, Z9, 543, 721; GSF6; G12

Table 1.2 Classification of anthraquinones as high-, medium- and low-active or no active NC inhibitors.

Since our hypothesis is that the mechanism of action could be through TAR and cTAR secondary structures stabilization, we analyzed the ability of all anthraquinones to interact and stabilize the sequences of cTAR, TAR and of the TAR/cTAR hybrid by Melting Assay.

Melting analysis demonstrates the affinity of the tested anthraquinones toward viral nucleic acid structures involved in minus strand transfer. The obtained results confirmed our hypothesis since the highly active compounds strongly stabilized the tested stem-loop structures. 2,6-dipeptidyl anthraquinones interact preferentially with bulge-loop structured TAR and cTAR rather than duplex structured hybrid TAR/cTAR. Anthraquinones prefer DNA sequences: this preference can be ascribed to the tighter nature of base stacking in canonical B-type helices than in either canonical A-type or hybrid structures [66].

All the anthraquinones of Z series result highly or medium active as inhibitors of NC activity; in particular Z9 is the most active both on TAR and on cTAR. Anthraquinones of Z series, moreover, recognize selectively TAR instead of non specific DNA [62]: this selective interaction could be important for therapeutic inhibition of NC-function.

Considering GSF series, the relationship between chemical structures of anthraquinones and their ability to interact with nucleic acids suggests that charges on lateral chains can play a significant role in recognition: a single charge is not enought to provide stabilization while two charges are necessary. Variation of charge distance from anthraquinonic nucleus results important for the nucleic acid structures stabilization: by looking at the structures of GSF6, GSF7 and GSF8 and their relative ΔTm data, the increasing charge distance from nucleus determines the higher ability of compounds to stabilize nucleic acid structures (ΔTm GSF6 > ΔTm GSF7 > ΔTm GSF8). Charge position on side chains turns out to be important also by comparing compounds GSF7 and GSF12. Both of them present two charges separated by the

same number of methylene groups: in GSF12 the charge is on the cycloesane substituent, while in GSF7 the charge is located at the end of the aliphatic residue. The flexibility of the chain, and hence of the charge position, determines an evident increasing in the nucleic acids stabilization. Also anthraquinones of GSF series recognize selectively TAR-RNA instead of non specific dsDNA [67]. From the correlation of NC inhibition and nucleic acids stabilization, GSF6 results to be the most promising anti-NC compounds of GSF series.

Series G compounds differ from GSF series for the side chains length and this modification determines a loss of anti-NC activity and also a low TAR and cTAR stabilization. The only promising compound is G12 that is highly active on cTAR and medium active on TAR.

We further analyzed our tested compounds on NC-assisted annealing of TAR with cTAR, to verify the effectiveness of the inhibition of the NC activity in steps subsequent to TAR and cTAR melting. We observed that an effective impairment of the NC-mediated melting of TAR and cTAR led to a corresponding reduction of the hybrid formation.

Hence, viral nucleic acids involved in reverse transcription represent a new validated target for anti-HIV drugs design and 2,6-dipeptydil anthraquinones were identified as highly active anti-NC compounds which could be further analyzed to be developed as new potential antiviral agents. The results obtained from this work could be considered a reliable starting point for further studies and for the design and for the synthesis of more active derivatives.

In this thesis viral nucleic acids were also investigated in the context of transctivation of HIV-1. Quinolones and in particular the 6-desfluoroquinolones (6-DFQs) are a promising class of anti-HIV compounds which interfere with the Tat mediated transcription process of the viral replicative cycle [47]. Structure–activity relationship (SAR) studies previously conducted [68-71] show that the highest antiretroviral activity was achieved when the essential 4-oxo-1,4-dihydroquinoline-3-carboxylic nucleus of these compounds was functionalized by a suitable 4-arylpiperazine substituent at the C7 position and an amino group or hydrogen atom at the C6 position. Although a methyl group emerged with optimal activity in the N1 position, the absence of a substituent at N1 also yielded acceptable anti-HIV activity.

In this work, new analogues that have different chemical modifications at the N1 position were analyzed by FQA, a FRET-based platform of screening performed and validated in our lab [50]. FQA allowed the determination of the inhibition constant Ki for each compound. We add new insights into the SAR of antiviral quinolones: all the tested 6-DFQs of WP series are active as Tat-TAR complex formation inhibitors, with Ki values lower than the control WM5.

Increasing the distance of the pyridin ring from the quinolone nucleus, the ability of the compounds to inhibit the formation of the complex peptide-TAR increases: Ki WP3-5 <WP6-5 <WP5-5; moreover, the presence of piperazine in the side chain (WP3-5 and WP6-5) has a positive effect on the activity of compounds, probably because it confers more rigidity to the chain and the position of the pyridine ring respect to the quinolone nucleus is fixed. The only compound of the HP series that has been possible to study is HP12-13N, which results to be active as inhibitors of the Tat-TAR complexation.

Both the WP and HP quinolones series resulted to be unstable in aqueous solution buffer over time; hence, their solubility has to be improved, to continue the development of quinolones as potential anti-HIV agents.

Ellagic acid (EA), a natural polyphenol, abundant in fruits and common in our diet, and some synthetic bi- and mono-lactones derivatives were also investigated, since, ellagic derivatives (ELs) showed antiviral activities: some of them inhibit the replication of Herpes Simplex virus by blocking the viral adsorption to cultured cells and others inhibit the reverse transcriptase activity in mouse leukemia virus infected cells [52]. ELs are under intense investigation in our lab not only as anti-topoisomerases but also as anti-NC agents: they were tested as NC inhibitors [65] and 1186 was the most active compound [65]. Moreover, Els resulted to be weakly cytotoxic on cell line MCF-7 and HeLa [72].

In this work ELs were analyzed to evaluate their inhibition activity on the Tat-TAR complex formation. ELs analysis was not possible to be performed by FQA due to a quenching effect observing after the addition of ELs to fluoresceinated Tat. The analysis was therefore performed by EMSA. Only compound 1186 can act as Tat-TAR complex formation inhibitor. 1186 is characterized by six hydroxyl groups that seem to have a crucial role in ELs activity. These results support the potential anti-HIV activity of ELs, and suggest the development of research on natural and synthetic compounds belonging to this class as potential anti-HIV agents.

2.

APTAMER BASED-MICROARRAY FOR PROTEIN DETECTION: A POTENTIAL TOOL FOR DIAGNOSTIC APPLICATIONS

2.1 INTRODUCTION

2.1.1 Aptamers

The conceptual framework and process of aptamer generation emerged from pioneering experiments by independent groups, both of whom published their work in 1990. Tuerk and Gold described a process of in vitro selection, dubbed "SELEX" (*S*ystematic *E*volution of *L*igands by *EX*ponential enrichment), to determine the sequence requirements of T4 DNA polymerase. An eight-base region of an RNA that interacts with T4 DNA polymerase was chosen and randomized. Two different sequences were selected from a calculated pool of more of 60000 species [73] SELEX identified the natural bacteriophage recognition sequence of the polymerase as the primary high-affinity ligand, along with a single major sequence variant that showed similar affinity for the enzyme. Shortly thereafter, Ellington and Szostak published further groundbreaking experiments utilizing in vitro selection to isolate nucleic acids with specific ligand binding properties [74]. Starting nucleic acid libraries were comprised of molecules containing 100 nucleotides of randomized sequence. Affinity chromatography was used to isolate RNAs that demonstrated highly specific binding to several organic dye molecules, none of which had been known previously to bind nucleic acid. Approximately 1 in 10^{10} random sequence RNAs was found to fold in a manner that created specific dye binding sites [73].

Ellington and Szostack originated the term "aptamer", derived from the Latin word "*aptus*" meaning "to fit", to describe these nucleic acid-based ligands. Since 1990, aptamers have been generated against hundreds of molecular targets, ranging from small molecules and peptides to numerous proteins of therapeutic interest (e.g., growth factors, enzymes, immunoglobulins, and receptors) [73].

Due to the development of SELEX, which is now a basic technique for the isolation of aptamers, many aptamers could be directly selected *in vitro* against various targets, from small biomolecules to proteins and even cells [75]. Aptamers are single stranded DNA or RNA typically from 15 to 40 nucleotides in length that bind to molecular targets with high affinity and specificity. The binding is based on hydrogen bonds, van der Waals and electrostatic interactions and depends on the exact sequences of nucleotides that comprise the aptamer: DNA or RNA ligand adopts a spatial conformation (folding), encoded in the primary sequence, making it selective for the target molecule. Aptamers are extraordinary tools: they bind with affinity, selectivity and specificity a wide range of target having *Kd* values in the

low nanomolar to picomolar range (Table 2.1) [76]. Moreover, aptamers can distinguish between closely related members of a protein family, or between different functional or conformational states of the same protein.

Target	K_D (nM)
Keratinocyte growth factor (KGF)	0.0003
HIV-1 reverse transcriptase (RT)	0.02
Transforming growth factor-beta 1 (TGF-β1)	0.03
P-selectin	0.04
Vascular endothelial growth factor receptor (VEGFr)	0.07
Platelet-derived growth factor (PDGF)	0.09
Immunoglobulin E (IgE)	0.1
Extracellular signal-regulated kinase (ERK)	0.2
CD4 antigen	0.5
HIV-1 RNAse H	0.5
Angiogenin	0.7
Complement factor C5	1.0
Transforming growth factor-beta 2 (TGF-β2)	1.0
Interferon-gamma (IFN-γ)	2.0
Secretory phospholipase A2 (sPLA2)	2.0
Thrombin	2.0
L-selectin	3.0
Human neutrophil elastase	5.0
Integrin ($\alpha_v\beta_3$)	8.0
HCV NS3 protease	10.0
Yersinia pestis tyrosine phosphatase	18.0

Table 2.1 Protein targets against which high affinity aptamers have been selected [76].

2.1.2 Aptamers *versus* Antibodies

The use of antibodies as the most popular class of molecules for molecular recognition in a wide range of applications has been around for more than three decades. Aptamers are widely known as a substitute for antibodies, because these molecules overcome the weaknesses of antibodies. The possible advantages of aptamers over antibodies are described below [77].

- ✓ *High stability of aptamers*. It is well known that proteins are easily denatured and lose their tertiary structure at high temperatures, while oligonucleotides are more thermally stable and maintain their structures over repeated cycles of denaturation/renaturation.

Hence, the greatest advantage of oligonucleotide-based aptamers over protein-based antibodies is their stability at elevated temperatures. Aptamers recover their native conformation and can bind to targets after re-annealing, whereas antibodies easily undergo irreversible denaturation [78]. Thus, aptamers can be used under a wide range of assay conditions.

- ✓ *Production of aptamers (synthesis/modification).* The identification and production of monoclonal antibodies are laborious and very expensive processes involving screening of a large number of colonies. Additionally, the clinical commercial success of antibodies has led to the need for very large-scale production in mammalian cell culture [79]. Moreover, immunoassays are required to confirm the activity of the antibodies in each new batch, because the performance of the same antibody tends to differ depending on the batch. However, aptamers, once selected, can be synthesized in quantity with great accuracy and reproducibility via chemical reactions. These chemical processes are more cost effective than the production of antibodies. Furthermore, aptamers can be easily modified by various chemical reactions to increase their stability and nuclease resistance [80]. Additionally, it is possible to introduce signal moieties, such as fluorophores and quenchers that greatly facilitate the fabrication of biosensors.

- ✓ *Low immunogenicity of aptamers*. Aptamers usually seem to be low-immunogenic and low-toxic molecules, because nucleic acids are not typically recognized by the human immune system as foreign agents. However, antibodies are significantly immunogenic, which precludes repeat dosing [81]. The Eyetech Study Group demonstrated that a VEGF-specific aptamer displayed little immunogenicity when given to monkeys in 1000-fold higher doses.

- ✓ *Variety of target.* In instances of toxins or molecules that do not elicit strong immune responses, it is difficult to identify and produce antibodies, but aptamers can be generated in sufficient numbers. Moreover, aptamers show a high affinity and specificity for some ligands that cannot be recognized by antibodies, such as ions or small molecules, indicating that employing aptamers as the recognition components may markedly broaden the applications of the corresponding biosensors [80].

Based on the many advantages described, aptamers are considered to be an alternative to antibodies in many biological applications.

Aptamers have been studied as a bio-material in numerous investigations concerning their use as a diagnostic and therapeutic tool and biosensing probe, and in the development of new drugs, drug delivery systems, *etc*. (Figure 2.1).

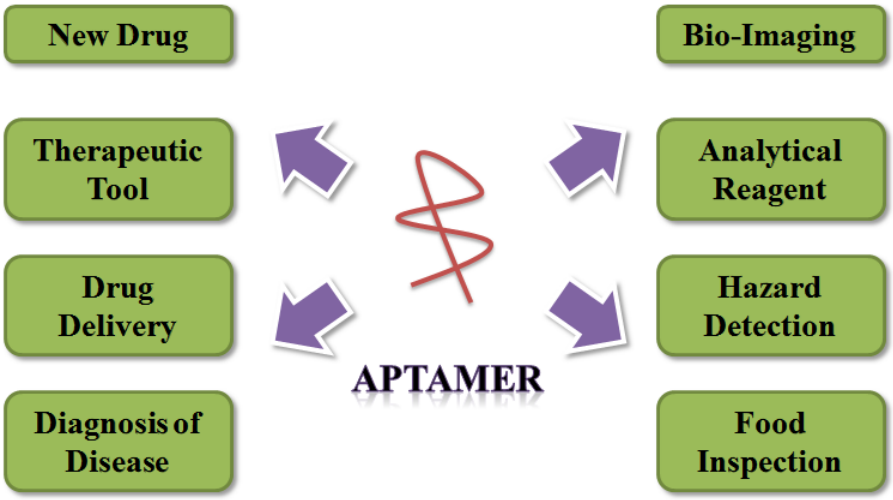


Figure 2.1 Various application of aptamers[77].

There have been attempts to search for aptamers that are specific to targets involved in various diseases, such as cancer and viral infection. Developed aptamers have been studied primarily for applications as diagnostic or therapeutic tools. In 2004, the approval by the Food and Drug Administration (FDA) of Macugen, a vascular endothelial growth factor (VEGF)-specific aptamer, for the treatment of neovascular (wet) age-related macular degeneration (AMD), is a prominent landmark in the application of aptamers [82]. Since then, aptamer technology has been regarded as more effective and more authoritative, and the number of studies on the applications of aptamers is rapidly increasing.

2.1.3 Aptamer-based Biosensors

Various definitions and terminologies are used depending on the field of application. Two commonly cited definitions by S.P.J. Higson and D.M. Frazer, respectively, are "a biosensor is a chemical sensing device in which a biologically derived recognition entity is coupled to a transducer, to allow the quantitative development of some complex biochemical parameter," and "a biosensor is an analytical device incorporating a deliberate and intimate combination of a specific biological element (that creates a recognition event) and a physical element (that transduces the recognition event)" [83]. A biosensor device is a combination of two parts: a bioelement and a sensor element. The basic concepts of a biosensor's operation is illustrated Figure 2.2. A specific bioelement recognizes a specific analyte and the sensor element transduces the change in the biomolecule into an electrical signal. The bioelement is very specific to the analyte to which it is sensitive. It does not recognize other analytes.

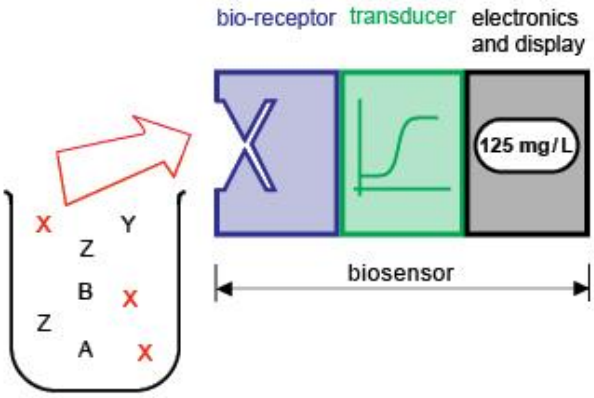


Figure 2.2 Biosensor principle. A biosensor consists of a bioreceptor for the specific detection of the respective analyte in spatial contact to a transducer for converting the signal into an electrically manageable format and a signal processing unit [84].

Depending on the transducing mechanism used, the biosensors can be of many types such as: resonant biosensors, optical-detection biosensors, thermal-detection biosensors, ion-sensitive field-effect transistor biosensors, and electrochemical biosensors. The electrochemical biosensors, based on the parameter measured, can be further classified as conductimetric, amperometric, and potentiometric. Typical bioreceptors in biosensors are enzymes, antibodies, microorganisms, and nucleic acids. Aptamers are a new promising group of bioreceptors, because of their outstanding selectivity, sensitivity and stability, the reproducibility of the target binding reaction, their production by chemical synthesis ensuring a constant lot-to-lot quality, and the ease of regeneration of aptamer derivated surfaces.

A biosensor that is based on aptamers as a recognition element is called an aptasensor. Aptasensors can be constructed through a variety of methodologies, including electrochemical biosensors, optical biosensors, and mass-sensitive biosensors, reported below [77].

Electrochemical Aptasensor

An electrochemical analysis is an attractive platform, because it offers high sensitivity, compatibility with novel microfabrication technologies, inherent miniaturization, and low cost. Therefore, various electrochemical aptasensors have been fabricated using several techniques. To enhance the sensitivity, an AuNRs- or AuNPs-modified conducting polymer was used as a material for immobilizing on the electrode [85]. Additionally, electroactive reporters, such as methylene blue (MB), ferrocene, ferrocene-bearing polymers, ruthenium complexes, and $Fe(CN)6^{4-/3-,}$ are used for signal transduction (Figure 2.3) [86]. In Gothelf's group, a signal-on electrochemical aptasensor was designed for theophylline detection using a redox-active Fc moiety, which could transfer [87] electrons with the electrode surface by DPV or CV [88].

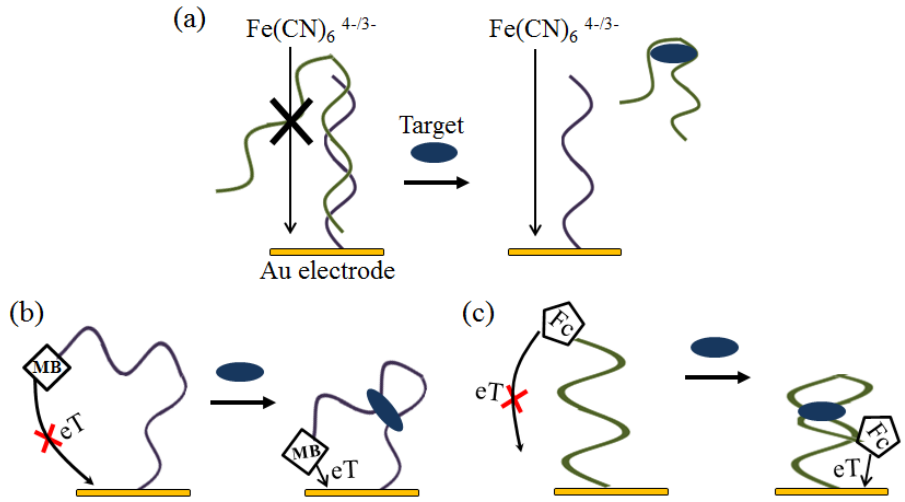


Figure 2.3 Examples of electrochemical aptasensors. (**a**) A schematic representation of the electrochemical aptasensor using $Fe(CN)6^{4-/3-}$. Part of the aptamer was a hybridized aptamer with complementary DNA, which was immobilized on the gold surface. In the presence of the target, the aptamer was followed by binding with the target, to decrease the amount of the aptamer on the electrode surface; (**b**) A schematic representation of the electrochemical aptasensor using MB. In the presence of the target, the aptamer folds into the target-binding three-way junction, altering the electron transfer (eT) and increasing the observed reduction peak; (**c**) A schematic representation of the electrochemical aptasensor using Fc. In the presence of the target, the aptamer folds into the restricted hairpin structure, and this conformational change results in increased efficiency of eT between the Fc probe and the electrode surface [77].

Fluorescence-Based Optical Aptasensor

Aptamers have been used as bio-probes in optical sensors based primarily on the incorporation of a fluorophore or a nanoparticle. In the case of fluorescence detection, the simplest format is to label the aptamers with both a quencher and a fluorophore. A cocaine-specific aptamer-based strategy was able to detect the target using a FRET signal between fluorescein and dabcyl moieties (Figure 2.4a) [89]. Aptamer beacons are probes that can monitor the presence of the target in solutions using fluorophore-quencher pairs. They can change into two or more conformations, such as hairpin shape or a hybridization form, by target binding. More complicated formats utilizing quaternary structural rearrangements that result in the assembly or disassembly of aptamers were also developed for fluorescence signaling (Figure 2.4b, c) [90]. Additionally, many nano-materials, including QDs, AuNPs, CNTs, graphene oxide (GO), polymer nanobelts, and coordination polymers, have been investigated for their fluorescence-quenching effect instead of using a more traditionally quencher [91, 92].

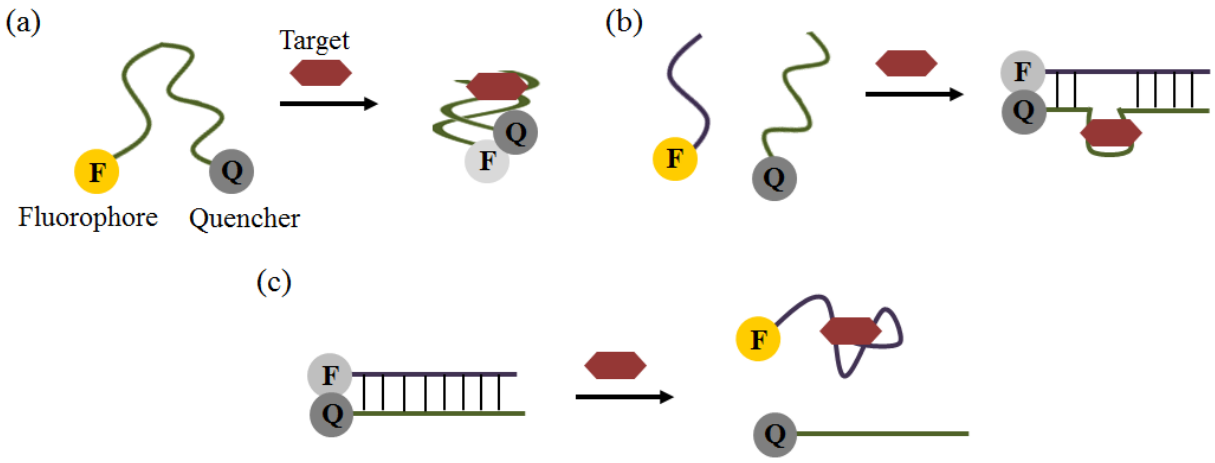


Figure 2.4 Schematic illustrations of optical aptasensors using fluorescence. (**a**) The simplest format of a quenching aptamer beacon. The binding of the target stabilizes the stem and brings the quencher and fluorophore in close proximity, resulting in fluorescence decrease; (**b**) Assembly aptamer beacon. The binding of the target brings the oligomers together and leads to ternary complex stabilization; (**c**) Disassembly aptamer beacon. The target binding induces an antisense displacement and results in a fluorescence increase [77].

Colorimetric-Based Optical Aptasensor

AuNPs or several polymers that cause color changes can be applied as novel reagents for the optical detection technique called colorimetry. The highly negatively-charged ssDNA (complementary strand of the aptamer), which is separated from the aptamer by interaction between the aptamer and the target, is stabilized against aggregation, and a color change occurs in conjunction with this phenomenon (Figure 2.5a) [93]. In contrast to this method, the

AuNP disaggregation method was employed to detect ATP, cocaine, Pb^{2+}, and K^{+} (Figure 2.5b) [94]. After hybridizing under the two kinds of oligonucleotides, the aptamer and a linker, the cross-linked AuNPs are released, because the aptamer undergoes a conformational change in the presence of the target. These colorimetric strategies have the virtue of being easily visible to the naked eye without the need for complicated instruments.

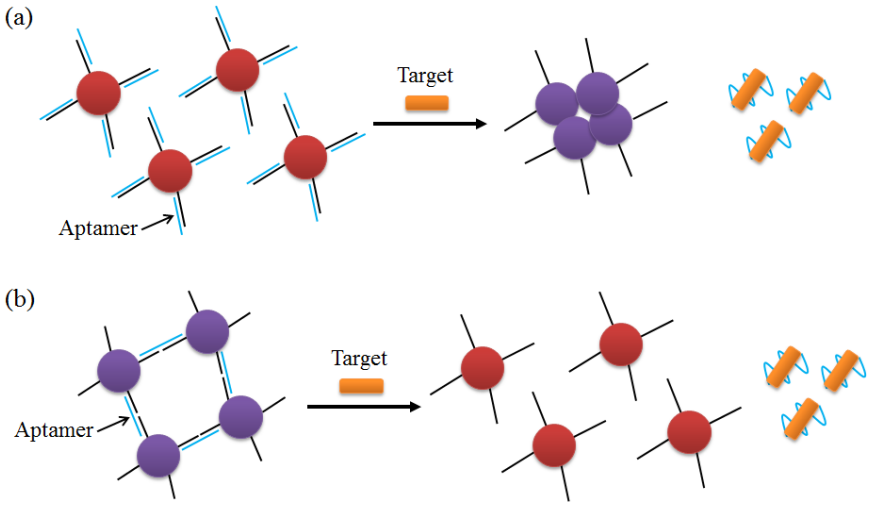


Figure 2.5 Schematic illustrations of optical aptasensors using AuNPs. (**a**) Aptamer release and AuNP aggregation by target binding; (**b**) Aptamer release and AuNP disaggregation by target binding [77].

Other Aptasensors

Numerous other aptasensors have been exploited in combination with various types of analytical equipment, such as those used for SPR, surface acoustic wave (SAW), QCM, and microchannel cantilever sensors (Figure 2.6) [95-97]. A nanoaptasensor using a single wall nanotube (SWNTs) device was developed for detecting small molecules [98]. These methods generally do not require target labeling; they only observe the differential changes in optical properties and mass. Therefore, these methods are primarily appropriate for relatively large molecules but not for small molecules, including organic molecules, toxins, and metabolites [99].

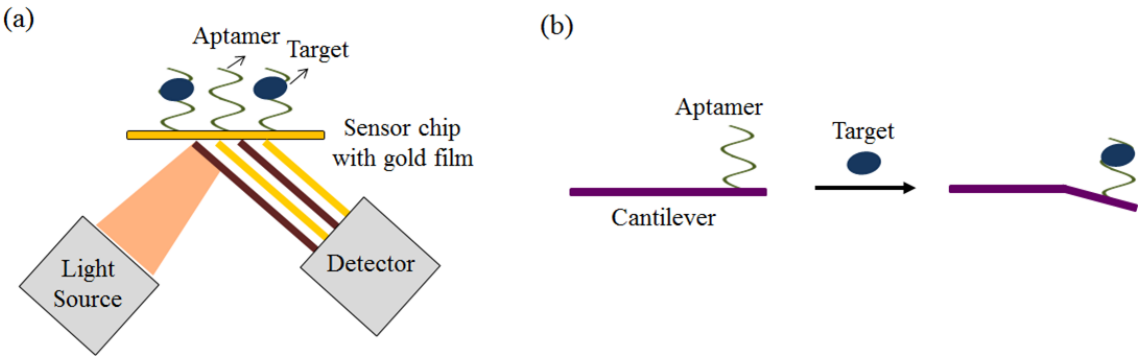


Figure 2.6 (a) SPR-based aptasensor and (b) microchannel cantilever-based aptasensor [77].

ELISA, one of the major clinical diagnostic tests available, is a versatile technique to detect almost any protein or peptide with high sensitivity. One version of ELISA, commonly referred to as sandwich ELISA, involves the simultaneous use of two antibodies or analyte-binding receptor proteins to capture the analyte or target and to report the target detection (Figure 2.7a). An aptamer-linked immobilized sorbent assay (ALISA) was introduced by Kiel's group. They demonstrated the feasibility of this method via a comparative study with ELISA using an antibody (Figure 2.7b). It is important to note that aptamers have an unlimited potential to circumvent the limitations associated with antibodies [100].

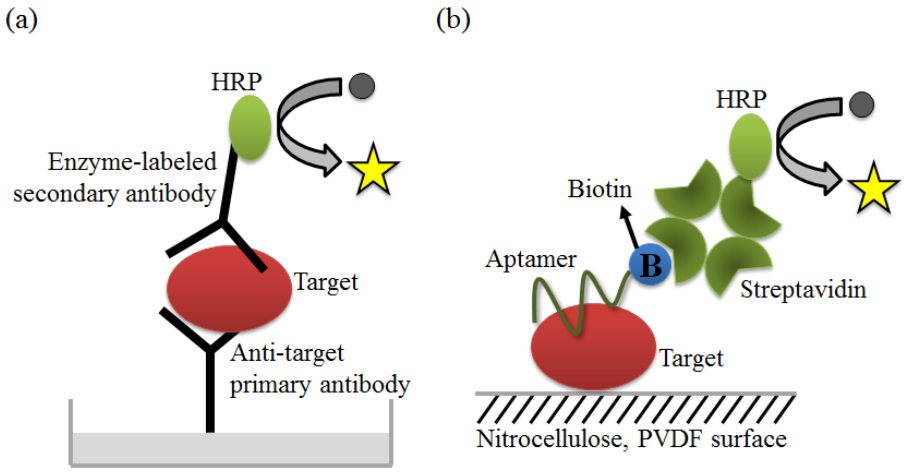


Figure 2.7 Schematic illustrations of (a) the ELISA method; and (b) the ALISA method [77].

In addition to ELISA, another common tool for clinical diagnosis is RDT, which is a rapid and simple method for point-of-care testing (POCT). RDTs for the diagnosis of infectious diseases, such as malaria and influenza, are already in use as commercially available tests. Because of the high sensitivity of the aptamer, several aptamer-based RDT methods, for many biomarkers related to diseases, have been introduced for use in early diagnosis or convenient POCT. In particular, Liu's group has developed a dry-reagent strip biosensor based on aptamers and functionalized AuNPs for use in thrombin analysis (Figure 2.8). In this study, the sensor is subject to visual detection of protein within minutes, with sensitivity and specificity that are superior to those of the antibody-based strip sensor [101]. The target interacts with the AuNP-primary aptamer conjugate, while the sample solution containing the target migrates onto each pad. Then, the target-aptamer-AuNP complexes are captured by the secondary aptamer that is immobilized in the test zone. Finally, a red band can be observed on the test zone due to the accumulation of AuNPs.

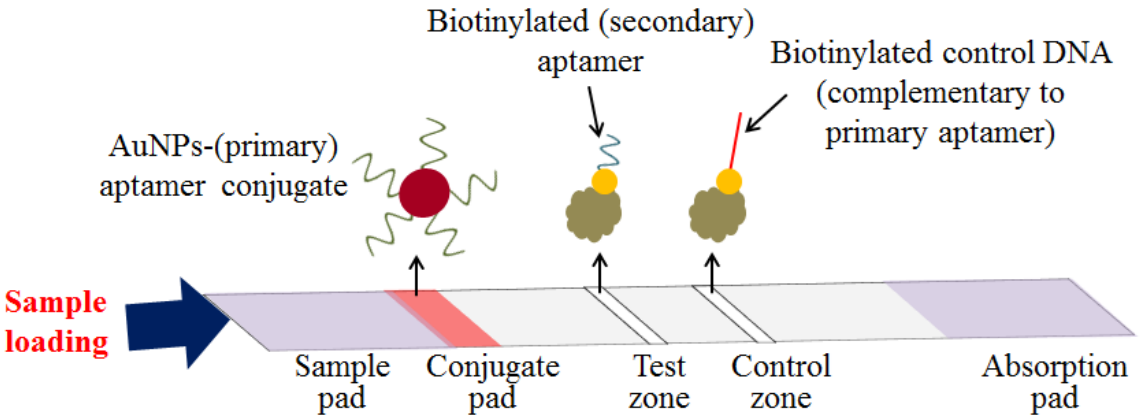


Figure 2.8 A schematic illustration of an AuNPs-based strip assay [77].

Biosensors are interesting tools offering certain operational advantages over standard photometric methods, notably with respect to rapidity, ease-of-use, cost, simplicity, portability, and ease of mass manufacture. Biosensors have been developed for more than 40 years now, and have been commercialized for some special applications like blood glucose and lactate measurement or bioprocess control, amongst others. However, they have not entered the market as much as expected, which is caused by several reasons. One reason is the instability of the biological recognition element of the biosensor (e.g. enzymes, cells or antibodies). Until now, proteins are detected mostly by antibodies in analytical formats like ELISA, immunobead assay, western blotting, microarrays and also biosensors [84]. Aptamers are equal to antibody concerning their binding affinities, but furthermore they provide decisive advantages, as before described.

In the following table (Table 2.2), aptamer biosensors for different protein targets are presented and listed according to the kind of nucleic acid of the aptamer (DNA or RNA), the transduction mode and their reporter units (mediators, enzymes, dyes, etc.). Also, the achieved detection limits and linear detection ranges are listed.

Target Protein	Aptamer	Type of Sensor, Reporter Unit	Detect. Limit, Linear Range
Thrombin	DNA beacon	ec, differential pulse voltammetry, methylene blue intercalator	11 nM 0 … 50.8 nM
Thrombin	DNA	ec, impedance spectroscopy, $[Fe(CN)_6]^{3-/4-}$	2 nM 5 … 35 nM
Thrombin	DNA thiolated/ biotinylated	ec, differential pulse polarography, p-nitroaniline/ peroxidase/ HRP	80 nM/ 3.5 nM n.s.
Thrombin	DNA labeled with methylene blue	ec, alternating current voltammetry, methylene blue	n.s. n.s. (logarithmic dependence)
Thrombin	DNA labeled with pyrroquinoline	ec, amperometry, glucose;	10 nM 40 ... 100 nM
Thrombin	DNA ferrocene labeled	optical combined with ec (cyclic voltammetry), eSPR/ ec, amperometry with co-immobilized microperoxidase	n.s. n.s.
Thrombin	DNA thiolated/ biotinylated	optical, SPR (Biacore™)	n.s. n.s.
Thrombin/ Lysozyme	n.s., thiolated	ec, square wave stripping voltammetry	0.5 pM (20 ... 500 ng/L)[1]
Lysozyme	DNA	ec impedance spectroscopy, $[Fe(CN)_6]^{3-/4-}$	
Lysozyme	DNA	ec, $[Ru(NH_3)_6]^{3+}$ cv peak decrease with target binding	0.5 µg/ml 0.5 … 50 µg/ml
IgE	DNA thiolated	optical, SPR	2 nM 8.4 ... 84 nM
IgE	DNA biotinylated	mass sensitive, QCM	100 µg/L 0 ... 10 mg/L
IgE	DNA	carbon nanotube FET	250 pM 250 pM … 20 nM
IgE	DNA	ec impedance spectroscopy, array	0.1 nM 2.5 … 100 nM
HIV-Tat protein	RNA biotinylated	optical, SPR/ mass sensitive, QCM	n.s./ 0.25 ppm 0 ... 2.5 ppm/ 0 ... 1.25 ppm
HIV-Tat 1 protein	RNA biotinylated	mass sensitive, QCM	0.65 ppm 0 ... 2.5 ppm
Abrin toxin	DNA	optical, luminescence, molecular light switching intercalator	1 nM 1 … 400 nM
Thrombin, bFGF, IMPDH, VEGF	RNA, DNA, fluorescently labeled	optical array, fluorescence polarization anisotropy	n.s.

ec electrochemical
n.s. not specified
[1] given as „analytically useful concentration dependence“

Table 2.2 Aptamer-based biosensors for protein detection [84].

2.1.4 Thrombin and Thrombin Binding Aptamers

Thrombin, a protein involved in the blood coagulation cascade, was the first biological macromolecule exploited for aptamer selection [102]. Thrombin is a serine protease that plays an important role in thrombosis and hemostasis. Thrombin is produced in the latest steps of the blood coagulation cascade. It converts fibrinogen into clottable fibrin; in addition, thrombin promotes platelet activation and aggregation [103, 104]. Thrombin consists of a 36-residue A-chain and a 259-residue B-chain [105], and is characterized by two different exosites, one binding fibrinogen (FBD) and the other binding heparin (HBD) (Figure 2.9).

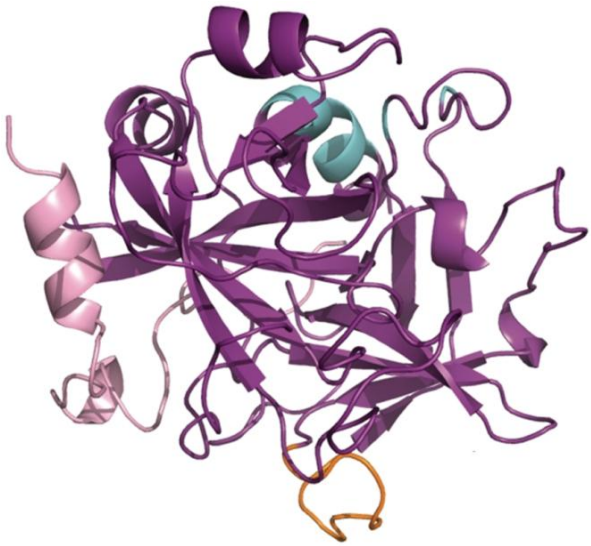


Figure 2.9 Thrombin molecule is represented as cartoon, with heavy chain colored purple, light chain colored light pink, exosite I colored orange and exosite II colored cyan; modified from [106].

Thrombosis is the formation of an improper blood clot (thrombus) in arteries or veins, due to an uncontrolled activation of the hemostatic system. Thrombosis reduces or obstructs the blood flow and/or could determine the detachment of thrombotic fragments that are able to occlude further downstream located vessels. Thrombotic disorders and their common clinical phenotypes of acute myocardial infarction (AMI), ischemic stroke, and venous thromboembolism cause substantial health care expenditures, morbidity, and mortality worldwide.

The concentration of thrombin in blood varies considerably, and can be almost absent in the blood of healthy subjects. However, it can reach low-micromolar concentrations during the coagulation process, and even low levels of thrombin can be generated in the early hemostatic process [107]. Thrombin, in addition to its direct actions on the coagulation system, has other functions as a potent signaling molecule that regulates physiologic and

pathogenic responses: it represents, for example, a potent chemotactic agent for monocytes and leukocytes involved in the atherosclerotic and thrombotic processes [108]; it is a potential pro-inflammatory mediator in neurotrauma and neurodegenerative disorders [109] and it plays a role in the development, plasticity and pathology of the nervous system [110]. Considering thrombin's physiological and pathological roles, it is not surprising that simple, low cost, and sensitive methods for its detection are actively pursued, and human thrombin is one of the preferred targets for the development of aptamer-based assays.

A variety of aptamer-based strategies designed for protein detection have been described in the past few years, employing in most cases two aptamers selected for targeting thrombin at its exosites: TBA1, a 15-nucleotide DNA aptamer able to bind FBD [111] and TBA2, a 29-nucleotide DNA aptamer able to recognize HBD [112].

TBA1 binds to FBD with a Kd value of 26 nM [113]. ^{1}H NMR spectroscopy on the 15mer TBA1 shows that the oligonucleotide forms a stable intramolecular G-4 structure comprising two G-quartets connected by two TT loops and one TGT loop [114] (Figure 2.10a). The structure of a complex between human α-thrombin and a 15-nucleotide aptamer has been solved by X-ray crystallography [115, 116]. Differences in the 15-mer loop conformations when in the complex or free in solution can be attributed to conformational changes occurring upon thrombin binding. Notably, no cations could be evidenced in the thrombin complex. Nonetheless, addition of cations, in particular of potassium, and/or thrombin tightens the G-4 structure. Hence, it appears that thrombin itself acts as a molecular chaperone, by promoting the quadruplex formation, independent of the presence of cations [117].

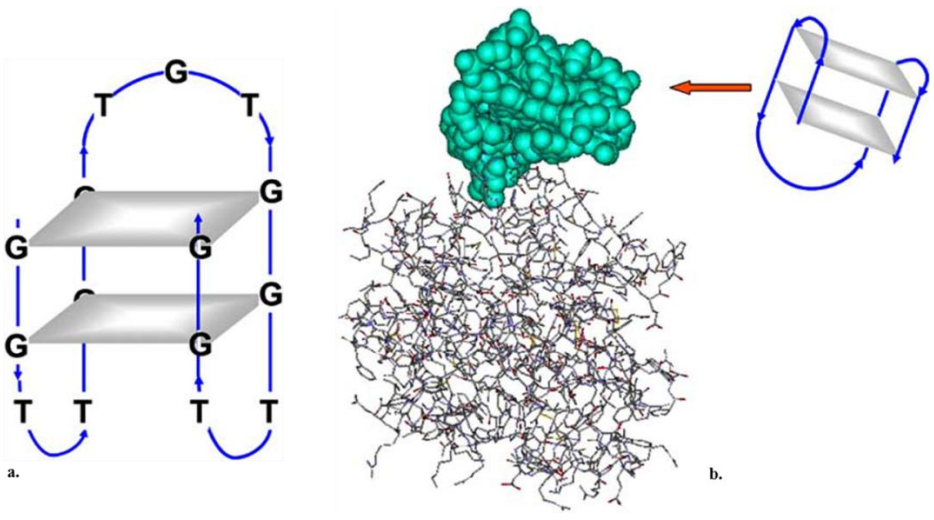


Figure 2.10 a. Secondary structure of the 15-mer thrombin binding aptamer TBA1; b. X-ray structure of TBA1 bound to α-thrombin; adapted from [116].

Subsequent studies led to the identification of a longer (29-nucleotide) aptamer with a substantially lower (0.5 nM) Kd value [112]. While TBA1 binds to the exosite I of thrombin, the TBA2 recognizes the heparin binding exosite. The likely conformation consists in a quadruplex/duplex with a 15-nucleotide "core" sequence that shares a close similarity with the previously described DNA ligand to thrombin, but binds 20 to 50-fold more tightly (Figure 2.11). Remarkably, a single T-A nucleotide exchange in the quadruplex loop is sufficient to direct the DNA to a distinct epitope. Besides confirming the importance of the G-4 structure, additional sequence information appears to reside in the duplex regions and contribute to greater stability and affinity in thrombin binding.

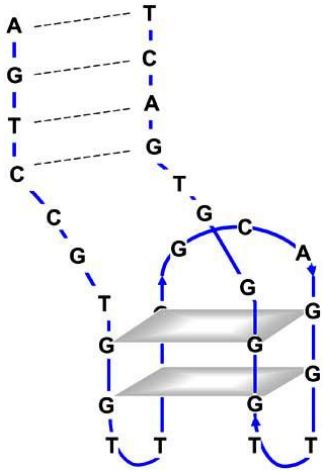


Figure 2.11 Secondary structure of the 29-mer thrombin binding aptamer TBA2 [116].

Aptamer systems described in literature that employ both TBA1 and TBA2 present different formats and detection strategies including fluorescence, electrochemistry, inductively coupled plasma mass spectrometry (ICP-MS) and surface plasmon resonance (SPR), yielding variable limits of detection (LOD), ranging from nanomolar [118-120] to picomolar [121-123].

Among the aptamers detection platforms described, microarray represents an appropriate sensing format for high-throughput analysis, allowing to analyze a great number of samples at the same time and to scale up the system in order to obtain a multi-sensing platform [124]. In addition, a twosite binding approach is suitable for high-sensitivity detection of a target protein with two or more distinct target domains [125], allowing the setup of a sandwich on a solid support: this approach ensures higher assay specificity because the detection takes place only if the analyte is simultaneously recognized by two different ligands. Thrombin and thrombin binding aptamers TBA1 and TBA2 can be used as model system (Figure 2.12).

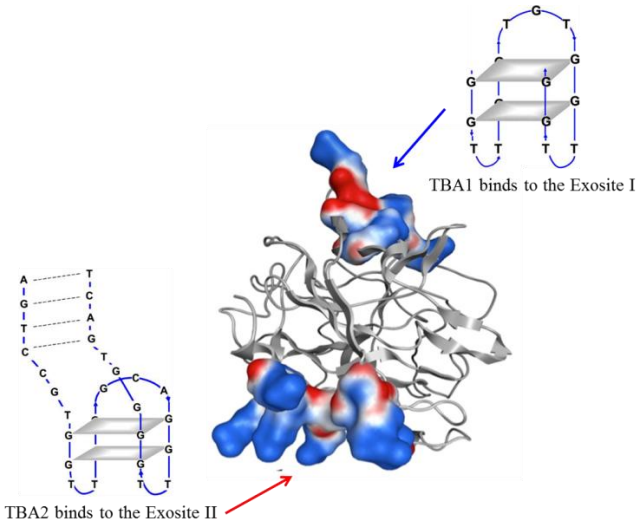


Figure 2.12 X-ray structure of human thrombin (pdb code: 1HUT). Surfaces corresponding to exosite I and II are evidenced. Surfaces charges are evidenced on the exosites surfaces (blue positive charges, red negative ones). The two non-overlapping aptamers recognize two different exosites of human thrombin [modified from [126].

2.1.5 $VEGF_{165}$ and Anti-VEGF Aptamers

Angiogenesis is a complex and highly regulated physiological process that consists in the sprouting of new capillaries from pre-existent vessels. Angiogenesis is involved in the embryonal development and in several physiological processes [127], and it is strictly controlled by over 30 different pro- and anti-angiogenic factors released in response to physical or chemical stimuli like hypoxia, presence of wounds or inflammation [128, 129]. The control of such a diverse pattern of mediators can undergo deregulation resulting in the formation of new vessels: neovascularization allows tumors to grow and to metastasize [8] and pathological angiogenesis play a crucial role also in non-neoplastic diseases as rheumatoid arthritis [130]. Another highly interesting area of investigation is represented by age-related retinopathies: wet Age-related Macular Degeneration (AMD) is the most common cause of irreversible vision loss. In all these diseases upregolation of a single pro-angiogenic factor, namely the Vascular Endothelial Growth Factor (VEGF), has been indicated as the link between inflammation and choroidal neovascularization [129]; excessive growth of blood vessels and increase of vascular permeability causes scarring of the eye and death of the photoreceptors in the retina leading to blindness.

Many antiangiogenesis treatment strategies currently in preclinical and clinical development focus on inhibition of the VEGF pathway. Two anti-VEGF biologics are

currently approved in Europe for the treatment of AMD: a monoclonal antibody (Mab) fragment (ranibizumab) and an aptamer (pegaptinib). Ranibizumab is the Fab fragment of a humanized Mab directed against human VEGF [131]. By intravitreal injection it finds applications in the treatment of AMD and macular oedema caused by diabetes. The anti-VEGF aptamer pagaptinib (Macugen®), approved for the treatment of AMD, binds to and inhibits the function of VEGF [131] present in the eye in meaningful levels and associated with blood vessel growth and leakage.

VEGF is a dimeric protein of variable molecular weight, found in several isoforms resulting from alternative splicing. $VEGF_{165}$ in particular is the predominant isoform sustaining pathological neovascularization. VEGF is a heparin-binding protein and its three dimensional structure has been determined [132-137] (Figure 2.13).

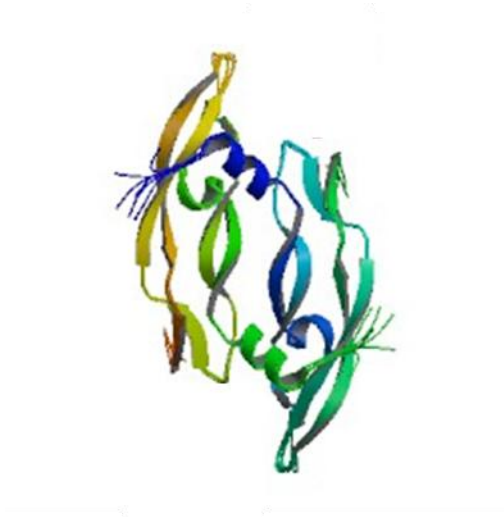


Figure 2.13 Three dimensional structure of VEGF.

$VEGF_{165}$ consists of two domains: the receptor-binding domain (RBD) and the heparin-binding domain (HBD) [138, 139]. VEGF represents an optimal target for the development of an aptasensor in a sandwich format: two DNA aptamers targeting $VEGF_{165}$ were recently selected and showed to exhibit different recognition properties.

VEa5 aptamer was selected in the presence of another protein used as competiror (PQQGDH) in order to isolate aptamers that have high specificity for $VEGF_{165}$ [140]. Measurement of the binding affinity was performed by surface plasmon resonance (SPR) and a Kd value of 130 nM was determined for VEa5 against $VEGF_{165}$. To identify the minimal sequence elements that confer high binding affinity to the target, the secondary structures of the aptamer was predicted using the Zuker DNA folding program: VEa5 forms three stem-

loop structures, and the shortest truncated aptamer has all three stem-loop structures, which means that all three are necessary for the binding with the target (Figure 2.14) [140].

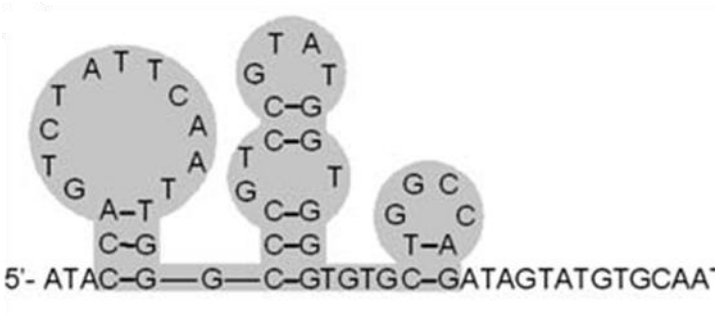


Figure 2.14 Predicted secondary structure and binding sites of VEa5. Secondary structure was predicted by m-fold. The portion required for binding to the target is shown in gray [140].

Furthermore, no binding was observed between VEa5 aptamer and the VEGF isoform $VEGF_{121}$ which lacks the HBD [138, 139]. This means that this aptamer may recognize the HBD of $VEGF_{165}$ [140].

An aptamer that binds to the RBD rather than the HBD of $VEGF_{165}$ was also developed. Three rounds of screening gave Vap7, which bound to the VEGF isoforms $VEGF_{121}$ and $VEGF_{165}$ with *K*d values of 1.0 nM and 20 nM, respectively. Moreover, Vap7 showed specificity within the VEGF family and did not bind to thrombin. Secondary structure predictions and circular dicrhoism suggested that Vap7 folds into a G-quadruplex structure in solution, well stabilized by potassium ions [141, 142].

Hence, two aptamers that bind to different regions of the target protein should permit the construction of a sandwich assay system. VEGF could be incorporated into sensor elements for cancer diagnosis. Despite its importance as a biomarker, very few sensing system employing aptamers for VEGF are described in the literature [143].

2.2 AIM OF THE WORK

This research project is aimed to build, optimize and characterize a novel protein/aptamer sensing system for the detection of proteins involved in diseases, with therapeutic and diagnostic interest. Biosensors based on the use of aptamers for specific recognition of an analyte are called "aptasensors".

We developed the Sandwich Aptamer Microarray (SAM, Figure 2.15) with the aim to replace the traditional immunochemical systems used as diagnostics. SAM is a system ELISA-like, using a similar architecture, but totally based on aptamer technology with a "primary aptamer" as selection element and a "secondary aptamer" as a signal transducer. Two DNA aptamers with distinct protein recognition patterns are used in tandem for the simultaneous recognition of the target protein. Appropriate chemical modifications were introduced in primary and secondary aptamers to adapt them to the format design and to the detection technology. To investigate whether these chemical modifications introduced in aptamers would affect target recognition, a deep analysis of the binary and ternary complexes formation in solution will be perform prior to the development of the SAM system in solid phase.

We aim to develop an innovative sensing system for the detection of human thrombin and of human $VEGF_{165}$ in biological samples to be used as a diagnostic device in clinical area.

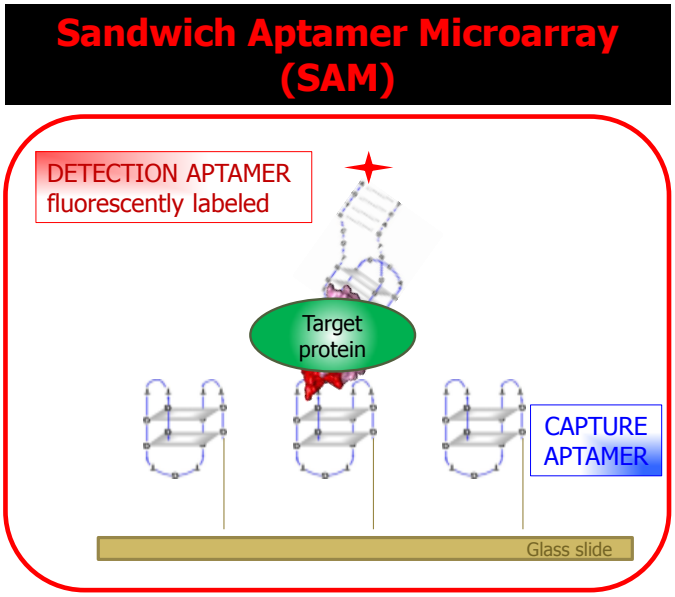


Figure 2.15 Representation of the Sandwich Aptamer Microarray (SAM). The "primary" aptamer, anchored on the glass slide, is used as capture layer for the target protein. Fluorescently labelled "secondary" aptamer is used as detection layer. Modification on the capture aptamer was the introduction of 5'-NH_2 group for aptamer anchorage on slide surface plus a polyT spacer for accessibility of aptamer to the target. Modification on the detection aptamer was the 5'-conjugation with the fluorescent label Cy5 for detection of the fluorescence signal by a laser scanner.

2.3 EXPERIMENTAL SECTION

2.3.1 Aptamers and Proteins

Thrombin Binding Aptamers

The sequence of the unmodified 15-mer thrombin binding aptamer TBA1 was: 5'-GGT TGG TGT GGT TGG-3'. TBA1 was used in its unmodified form as control; to allow immobilization on solid phase, a 5'-amino modification (TBA1-NH_2) and a 5'-amino modification plus a polyT(12) tail as spacer [reported as TBA1(12T)NH_2] were also synthesized.

TBA1: 5'-GGT TGG TGT GGT TGG-3';
TBA1-NH_2: 5'-NH_2-GGT TGG TGT GGT TGG-3';
TBA1(12T)NH_2: 5'-NH_2-TTT TTT TTT TTT GGT TGG TGT GGT TGG-3'.

The sequence of the unmodified 29-mer thrombin binding aptamer TBA2 control (TBA2) was: 5'-AGT CCG TGG TAG GGC AGG TTG GGG TGA CT-3'. TBA2 was used as 5'-Alexa Fluor 488 labeled (TBA2-Alexa) and as 5'-Cy5 labeled (TBA2Cy5) to allow detection. TBA2 was used also with a 5'-biotin modification (TBA2-biotin) to allow an indirect method of detection.

TBA2: 5'-AGT CCG TGG TAG GGC AGG TTG GGG TGA CT-3';
TBA2-Alexa: 5'-Alexa488-AGT CCG TGG TAG GGC AGG TTG GGG TGA CT-3';
TBA2Cy5: 5'-Cy5-AGT CCG TGG TAG GGC AGG TTG GGG TGA CT-3';
TBA2-biotin: 5'-biotin-AGT CCG TGG TAG GGC AGG TTG GGG TGA CT-3'.

In Figure 2.16 are reported the chemical modifications introduced to TBA2.

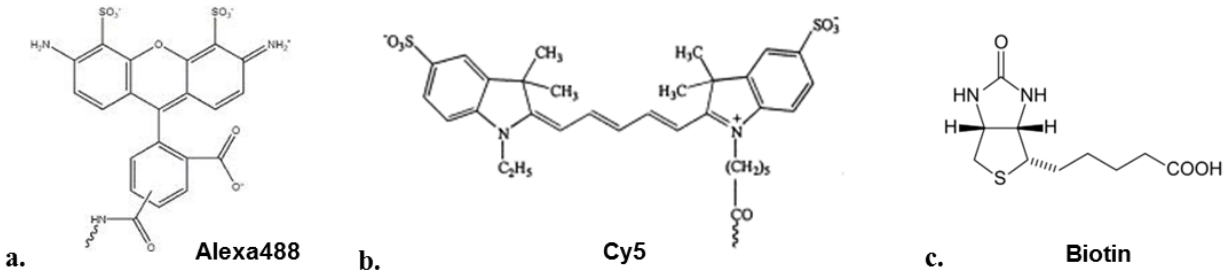


Figure 2.16 Chemical structures of Alexa Fluor 488 (a) and Cy5 (b) dyes. Alexa Fluor 488 dye is a green light emitter fluorophore with the following spectral properties: Absorption max at 495nm, Emission max at 519nm. Cy5 is a red light emitter fluorophore with Absorption max at 649nm and Emission max at 670nm. **Chemical structure of biotin (c).** Biotin binds very tightly to the tetrameric protein Streptavidin.

As negative control in the Sandwich Aptamer Microarray we used an aptamer (OTA) that binds with high affinity and specificity to ochratoxin A [144]; the sequence of 5'-amino modified OTA to allow the binding on the slide was: 5'-NH_2-GAT CGG GTG TGG GTG GCG TAA AGG GAG CAT CGG ACA-3'.

All the aptamers were purchased from IDT Integrated DNA Technologies (Munich, Germany) and stored at −20 °C in TE (10mM Tris-HCl, 1mM EDTA) pH 8.0.

Aptamers against $VEGF_{165}$

The DNA sequences of the unmodified aptamers against $VEGF_{165}$ were:

Vap7: 5'-ATA CCA GTC TAT TCA ATT GCA CTC TGT GGG GGT GGA CGG GCC GGG TAG A-3';
VEa5: 5'-ATA CCA GTC TAT TCA ATT GGG CCC GTC CGT ATG GTG GGT GTG CTG GCC AGA TAG TAT GTG CAA TCA-3'.

To allow their immobilization on solid phase, both the 49-mer Vap7 aptamer and the 66-mer VEa5 aptamer were synthetized with a 5'-amino modification (Vap7-NH_2 and VEa5-NH_2). Vap7 was also synthetized with a 5'-amino modification plus a polyT(12) tail as spacer [reported as Vap7(12T)NH_2].

Vap7-NH_2: 5'-NH_2- ATA CCA GTC TAT TCA ATT GCA CTC TGT GGG GGT GGA CGG GCC GGG TAG A-3';
Vap7(12T)NH_2: 5'-NH_2- TTT TTT TTT TTT ATA CCA GTC TAT TCA ATT GCA CTC TGT GGG GGT GGA CGG GCC GGG TAG A-3';
VEa5-NH_2: 5'-NH_2- ATA CCA GTC TAT TCA ATT GGG CCC GTC CGT ATG GTG GGT GTG CTG GCC AGA TAG TAT GTG CAA TCA-3'.

Both Vap7 and VEa5 were used as 5'-FAM labeled (Vap7-FAM and VEa5-FAM) and VEa5 aptamer was used also as 5'-Cy5 labeled (VEa5-Cy5) to allow detection:

Vap7-FAM: 5'-FAM- ATA CCA GTC TAT TCA ATT GCA CTC TGT GGG GGT GGA CGG GCC GGG TAG A-3';

VEa5-FAM: 5'-FAM- ATA CCA GTC TAT TCA ATT GGG CCC GTC CGT ATG GTG GGT GTG CTG GCC AGA TAG TAT GTG CAA TCA-3';
VEa5-Cy5: 5'-Cy5- ATA CCA GTC TAT TCA ATT GGG CCC GTC CGT ATG GTG GGT GTG CTG GCC AGA TAG TAT GTG CAA TCA-3'.

The chemical structure of the fluorophore 6-carboxyfluorescein is reported in Figure 2.17; the Cy5 chemical modification introduced to VEa5 is reported in Figure 2.16b.

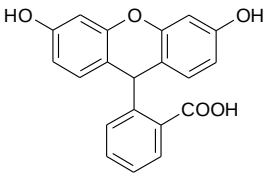


Figure 2.17 Chemical structure of the fluorophore 6-carboxyfluorescein.

All the aptamers were purchased from IDT Integrated DNA Technologies (Munich, Germany) and stored at −20 °C in TE (10mM Tris-HCl, 1mM EDTA) pH 8.0.

Thrombin and Protein Labeling
All protein molecules were purchased from Sigma-Aldrich (St. Louis, MO, USA).
Human α-thrombin (T1063) was dissolved in deionized purified (MilliQ) water, aliquoted and stored as recommended by the supplier; Bovine Serum Albumin (BSA), Lysozyme and Human Vascular Endothelial Growth Factor 165 (VEGF) were used to evaluate aptamer microarray specificity, and Cy3-labeled Streptavidin was used to perform thrombin assay in the presence of TBA2-biotin aptamer.
Fetal Bovine Serum (FBS) (Gibco, Monza, Italy) was used in order to obtain a complex sample matrix.

The AlexaFluor®555 monoreactive Succinimidyl Esters (Invitrogen), dissolved in anhydrous dimethyl sulfoxide (DMSO) solvent was used to label thrombin samples. Prior to labeling, thrombin and AlexaFluor® 555 were first diluted to proper reaction concentrations in MilliQ H_2O. The labeling reaction was done mixing thrombin samples with the AlexaFluor®555 NHS esters reagent in the Coupling Buffer (sodium carbonate 0.1 M, pH 9.4); the optimal dye to protein ratio was 2; the reaction was incubated 2 h at 4°C. After this time, glycine was added to the mixture in order to block the free unreacted AlexaFluor®555 NHS esters. The

same protocol was used also to label thrombin with AlexaFluor®647. The labeled protein samples were kept in the dark at 4°C, and all samples were used within a week.

$VEGF_{165}$ and Protein Labeling

Vascular Endothelial Growth Factor 165 ($VEGF_{165}$) human recombinant, expressed in *E.coli* was purchased from Sigma-Aldrich (St. Louis, MO, USA). Human $VEGF_{165}$ was dissolved in TBSE buffer (EDTA 0.05 mM, Tris-HCl 10mM, NaCl 100mM, pH 7) in order to obtain a 1 mg/mL stock solution as recommended by the supplier, aliquoted and stored at -20°C.

The AlexaFluor®555 monoreactive Succinimidyl Esters (Invitrogen), dissolved in anhydrous dimethyl sulfoxide (DMSO) solvent was used to label $VEGF_{165}$ samples. Prior to labeling, the protein and AlexaFluor® 555 were first diluted to proper reaction concentrations in MilliQ H_2O. The labeling reaction was done mixing $VEGF_{165}$ samples with the AlexaFluor®555 NHS esters reagent in the Coupling Buffer (sodium carbonate 0.1 M, pH 9.4); in this case, the optimal dye to protein ratio was 1; the reaction was incubated 2 h at 4°C. After this time, glycine was added to the mixture in order to block the free unreacted AlexaFluor®555 NHS esters. The labeled protein samples were kept in the dark at 4°C, and all samples were used within a week.

2.3.2 Analysis in Solution

2.3.2.1 EMSA (Electrophoresis Mobility Shift Assay) Analysis for Binary Complexes

Thrombin and thrombin binding aptamers

The binding of each aptamer to the protein target in solution was analyzed by EMSA. Prior to incubation with thrombin, all the aptamer samples (oligonucleotide 10 μM in the presence of KCl 100 mM) were denatured at 95 °C for 5 minutes and the samples were left to cool to room temperature (slow annealing). Each folded aptamer (1 μM) was then incubated with increasing concentrations (from 0 to 12 μM) of human thrombin in a total volume of 20 μL, at 25 °C for 30 min. After incubation, free aptamer and thrombin-aptamer complexes were resolved by 12% non-denaturing polyacrylamide gels containing 1X TBE buffer (Tris-HCl 89 mM, borate 89 mM, EDTA 2mM) and KCl 10 mM. Aptamers on the gels were stained with the fluorescent DNA binding dye SybrGreen II® (Invitrogen) that preferentially binds to single-stranded DNA and emits fluorescence (λ_{exc} = 488 nm; λ_{em} = 522 nm) when in

complex with DNA. Fluorescence in gel systems was detected on a Geliance 600 Imaging System (PerkinElmer).

$VEGF_{165}$ and anti-VEGF aptamers

The binding of each aptamer to the protein target in solution was analyzed by EMSA. Prior to incubation with $VEGF_{165}$, all the aptamer samples were folded: unmodified and modified VEa5 aptamers were folded in TBSE buffer (EDTA 0.05 mM, Tris-HCl 10mM, NaCl 100mM, pH 7), while to fold unmodified and modified Vap7 aptamers is necessary the presence of KCl and therefore TBSE buffer was added of 100 mM KCl. Vap7 aptamers 1 μM in the presence of TBSE buffer and KCl 100 mM or VEa5 aptamers 1 μM in TBSE buffer were denatured at 95 °C for 5 minutes and the samples were left to cool to room temperature (slow annealing). Each folded aptamer (0.1 μM) was then incubated with increasing concentrations (from 0 to 10 μM) of $VEGF_{165}$ in a total volume of 10 μL, at 25 °C for one hour. After incubation, free aptamer and protein-aptamer complexes were resolved by 12% non-denaturing polyacrylamide gels containing Glycine buffer 50 mM pH 9.5 and KCl 10 mM (Vap7 aptamers) or Glycine buffer 50 mM pH 9.5 (VEa5 aptamers). In the case of $VEGF_{165}$ is very important the choice of the buffer in the gels: a buffer with a pH>pI of the protein is needed to have the protein negatively charged and in this way the protein can enter into the gel systems. $VEGF_{165}$ has a pI of 8.5 and therefore a strongly basic buffer is required: glycine 50 mM pH 9.5 (with or without KCl) was used as buffer in the gels and as running buffer for EMSA.

Different methods were used to visualize oligonucleotides in gel systems: Vap7, Vap7-NH_2, VEa5 and VEa5-NH_2 aptamers on the gels were stained with the fluorescent DNA binding dye SybrGreen II® (Invitrogen) that preferentially binds to single-stranded DNA and emits fluorescence (λ_{exc} = 488 nm; λ_{em} = 522 nm) when in complex with DNA; the fluorescence of Vap7-FAM and VEa5-FAM in gel systems was directly detected on a Geliance 600 Imaging System (PerkinElmer). Moreover, Vap7 was labeled using T4 polynucleotyde kinase (Ambion, Applied Biosystems, TX, USA) in the presence of [γ^{32}P] ATP (3000Ci/mmol, Perkin Elmer, MA, USA) following manufacturer's instructions. EMSA of the [γ^{32}P] ATP radiolabelled Vap7 (Vap7*) aptamer was performed in the same experimental conditions explained above. Dried gel was put in contact with a sheet of europium for 24 hours. The sheet, acquired the radioactive signal, was scanned using al Phosphorimager STORM B40 GE Healthcare, Italy.

2.3.2.2 Supershift Assays for Ternary Complexes

TBA1-Thr-TBA2 ternary complexes

To verify the sandwich formation in solution a Supershift Assay was performed: both pairs of folded aptamers (each 1 μM) were incubated simultaneously with the protein (5 μM) in the Binding Buffer (Tris 20 mM, KCl 5 mM, NaCl 140 mM, MgCl2 1 mM, pH 7.5) in a total volume of 20 μL, at 25 °C for 30 min. After incubation, free aptamers and thrombin-aptamers complexes were resolved by 12% non-denaturing polyacrylamide gels containing 1X TBE buffer (Tris-HCl 89 mM, borate 89 mM, EDTA 2mM) and KCl 10 mM. Aptamers on the gels were stained also in this case with the fluorescent DNA binding dye SybrGreen II® (Invitrogen) that preferentially binds to single-stranded DNA and emits fluorescence (λ_{exc} = 488 nm; λ_{em} = 522 nm) when in complex with DNA. Fluorescence in gel systems was detected on a Geliance 600 Imaging System (Perkin Elmer). Relative mobilities of each band were calculated as Rf, *i.e.*, the ratio between the distance traveled by the sample (band) divided by the distance traveled by the marking dye (bromophenol blue) in the gel.

Vap7-VEGF-VEa5 ternary complexes

To verify the sandwich formation in solution a Supershift Assay was performed and two possible combinations of aptamers were tested: Vap7-NH_2 as capture layer and VEa5-FAM as detection layer; VEa5-NH_2 as capture aptamer and Vap7-FAM as detection aptamer. Both pairs of folded aptamers (each 0.1 μM) were incubated simultaneously with the protein (5 and 10 μM) in TBSE buffer (EDTA 0.05 mM, Tris-HCl 10mM, NaCl 100mM, pH 7) in a total volume of 10 μL, at 25 °C for 30 min. After incubation, free aptamers and $VEGF_{165}$-aptamers complexes were resolved by 12% non-denaturing polyacrylamide gels containing Glycine buffer 50 mM pH 9.5 and KCl 10 mM. Aptamers on the gels were stained also in this case with the fluorescent DNA binding dye SybrGreen II® (Invitrogen) that preferentially binds to single-stranded DNA and emits fluorescence (λ_{exc} = 488 nm; λ_{em} = 522 nm) when in complex with DNA. Fluorescence in gel systems was detected on a Geliance 600 Imaging System (Perkin Elmer).

2.3.3 Analysis in solid phase

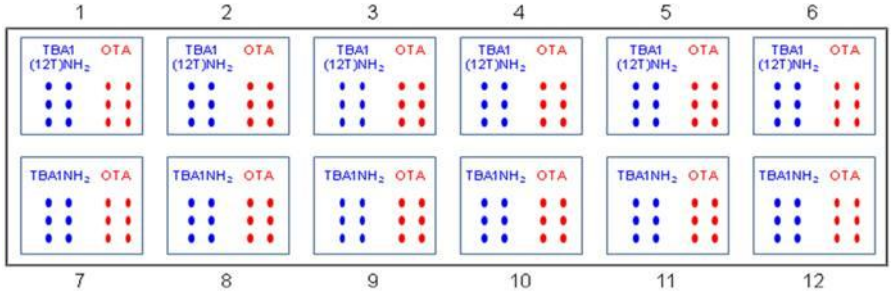
Aptamer Preparation for SAM

In the aptamer arrays TBA1 was used as capture layer for thrombin and Vap7 as capture aptamer for $VEGF_{165}$. A sequence that binds with high affinity and specificity to ochratoxin A, was anchored as negative control on a distinct spot on the slide surface parallel to TBA1's spots. Prior to immobilization on the glass slide, TBA1 and OTA (80 μM in presence of KCl 100 mM) and Vap7 (80 μM in TBSE buffer and KCl 100 mM) were denatured at 95 °C for 5 minutes and then left to cool to room temperature. This folding step ensures that they assume the G-quadruplex structure, responsible for protein binding. Folded aptamers were then diluted in the Printing Buffer 1.5X (Printing Buffer 6X: sodium phosphate300 mM, 0.02% Triton pH 8.5) to a final concentration of 20 μM. These probes were loaded into micro-plates and submitted to the Spotter Arrayer (Versarray Chip Writer Pro System, BioRad) for slide printing.

Printing of Aptamer Microarrays

E-surf LifeLine slides (25 mm × 75 mm, LifeLineLab, Pomezia, Italy) were used for printing: these slides are obtained by adsorption on glass of a hydrophilic polymer containing N,N-acryloyloxysuccinimide (NAS), allowing binding to the 5'-amino modified DNA aptamers TBA1-NH_2, OTA and Vap7-NH_2. To improve capture-protein efficiency, we anchored on the glass slide also the TBA1 and the Vap7 5'-amino modified with a polyT (12T) spacer at the 5'-end. In this way the G-quadruplex structure of TBA1 and Vap7 can fold correctly to recognize the protein target. The same spot scheme shown in Table 2.3 was printed 12 times (12 sub-arrays) on the same slide: thrombin capture aptamers [TBA1-NH_2 or TBA1(12T)NH_2] and OTA control (Table 2.3a) and capture aptamers against $VEGF_{165}$ [Vap7-NH_2 or Vap7(12T)NH_2] and OTA control (Table 2.3b) were printed in six spots to ensure the reproducibility of the system.

a.



b.

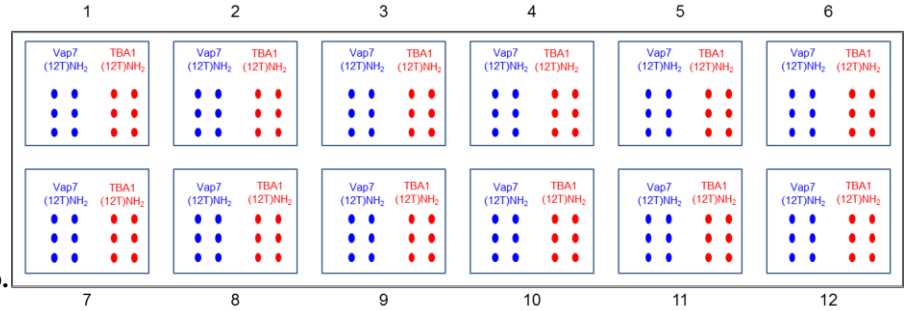


Table 2.3 Scheme of the printed glass slide for thrombin detection (a) and for $VEGF_{165}$ detection (b).

Slides were printed by the Spotter Arrayer instrument, using Telechem SMP3 microspotting pins. Printed slides were incubated overnight in a 75% humidity incubation chamber, blocked at 30 °C in Blocking Solution (0.1 M Tris, 50 mM ethanolamine), washed at 30 °C in Washing Solution (4X SSC, 0.1% SDS) and finally washed in MilliQ H2O, spin-dried and stored properly until usage. Each slide has the possibility to test up to 12 samples at once, since each sub-array can be physically isolated from the others by the Corning Microarray Hybridization Chamber (Corning, NY, USA) during the hybridization phase of the experiment. This approach ensures to hybridize up to 12 samples at the same time and nevertheless to minimize array variation resulting from minor fluctuation of external parameters.

Detection Layer Preparation

The fluorescently labeled TBA2 (TBA2-Alexa, TBA2Cy5 or TBA2-biotin) was used as secondary aptamer for detecting thrombin captured by the primary aptamer (TBA1). TBA2 (10 µM in KCl 100 mM) was denatured at 95 °C for 5 minutes and then left to cool to room temperature in order to assume G-quadruplex structure, which is essential for recognition of the heparin binding domain of thrombin.

The VEa5-Cy5 aptamer was used as detection aptamer against $VEGF_{165}$: VEa5-Cy5 (10 µM in TBSE buffer) was denatured at 95 °C for 5 minutes and then left to cool to room temperature in order to assume the correct structure for recognition of the heparin binding domain of $VEGF_{165}$.

Sandwich Aptamer Microarray Assays

The printed aptamer microarrays prepared as detailed above were immersed just before analysis in the Binding Buffer 1X (Tris 20 mM, KCl 5 mM, NaCl 140 mM, MgCl2 1 mM, pH

7.5) in the case of thrombin or in TBSE (EDTA 0.05 mM, Tris-HCl 10mM, NaCl 100mM, pH 7) containing KCl 100 mM in the case of $VEGF_{165}$, at room temperature for 15 min.
First, the SAM assays were performed using the thrombin system as model system in order to set up the best procedure and improve the protocol. The sandwich aptamer microarrays were performed with two procedures, *i.e.*, analyzing the sandwich formation on surface either in one step or in a two steps protocol (Figure 2.18).

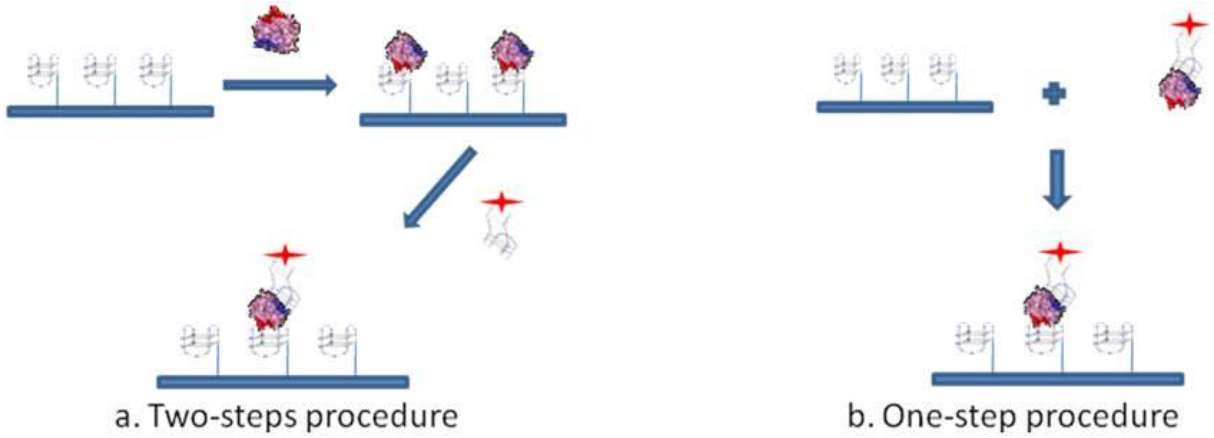


Figure 2.18 Representation of the Sandwich Aptamer Microarray (SAM) for thrombin detection performed following the two-steps (a) and the one-step (b) procedure. TBA1, anchored on the glass slide, is used as capture layer for thrombin. Fluorescently labeled TBA2 is used as detection layer. Modification on TBA1 was the introduction of 5'-NH_2 group for aptamer anchorage on slide surface plus a polyT spacer for accessibility of the G-quadruplex to thrombin. Modification on TBA2 was the 5'-conjugation with the fluorescent label Cy5 for detection of the fluorescence signal by a laser scanner [145].

The one step-procedure consisted in the pre-incubation of thrombin with the fluorescently labelled aptamer (TBA2Cy5) in solution in the Binding Buffer, at 25 °C for 30 min. The pre-formed complex thus obtained was then incubated in the microarray at 25 °C for 30 min. The two steps-procedure consists of the incubation in the microarray of thrombin diluted in the Binding Buffer at 25 °C for 30 min, followed by the incubation with the fluorescently labelled secondary aptamer (TBA2Cy5) diluted in the Binding Buffer, at 25 °C for further 30 min.
Thrombin samples were pre-incubated with different detection aptamer (TBA2-Alexa, TBA2-Cy5 or TBA2-biotin) in solution with thrombin Binding Buffer 1X, at 25 °C for 30 min. The pre-formed complexes thus obtained were then incubated on the microarray at 25 °C for 2 hours in the case of TBA2-Alexa and TBA2-Cy5. Steptavidin-Cy3 at the final concentration of 30 nM was added after 1h of incubation in the case of TBA2-biotin, and the system allowed to incubating for 1 additional hour.

Also the SAM assays for $VEGF_{165}$ detection were performed following both the one-step procedure and the two-steps procedure, using VEa5-Cy5 as detection aptamer and TBSE as binding buffer.

Finally, the aptamer microarrays were carefully rinsed with the PBS Buffer (phosphate 10 mM, NaCl 137 mM, KCl 3 mM, pH 7.4), washed three times with PBS at room temperature to remove the unbound proteins and rapidly rinsed in MilliQ H_2O before spin-drying, ready to be scanned.

Slides Scanning and Data Analysis

Spin-dried aptamer arrays were scanned using Genepix 4000B laser scanner (Molecular Devices) and the Gene Pix Pro software using both 532 nm and 635 nm wavelength and PTM gain 550 for both channels (Power 33). Fluorescent spot intensities were quantified using the Gene Pix Pro software after normalizing the data by subtracting local background from the recorded spot intensities. For each set of six spots median and standard deviation were calculated.

2.4 RESULTS AND DISCUSSION

2.4.1 Human thrombin detection through a Sandwich Aptamer Microarray

2.4.1.1 Analysis in Solution

2.4.1.1.1 Analysis and optimization of TBA1-Thrombin interaction

Role of K^+ Ions

TBA1 is a 15-mer DNA thrombin binding aptamer that recognizes specifically the Fibrinogen Binding Domain (FBD) of thrombin by adopting the four-stranded DNA structure known as G-quadruplex (G-4) [111, 114]. Monovalent cations, in particular K^+ and Na^+, are essential to form and stabilize G-4 structures [146], with K^+ ions favoring the recognition of thrombin to immobilized aptamers at temperatures higher than 25 °C [147]. The folding of TBA1 into G-quartets in solution is favored by the presence of K^+ ions, being 5.8 mM the minimal concentrations required for TBA1, while TBA2 requires a slightly higher ion concentration [148]. Therefore, we investigated the effect of potassium ions 10 mM on the binding of TBA1 and thrombin in our conditions to decide whether it would be a component of the buffer system. TBA1 was denatured, folded and incubated with thrombin (up to 12 μM) in the presence or absence of KCl. Complex formation was analyzed by EMSA. Free and complexed aptamers were detected in the gel by staining with SybrGreen II. The result is shown in Figure 2.19.

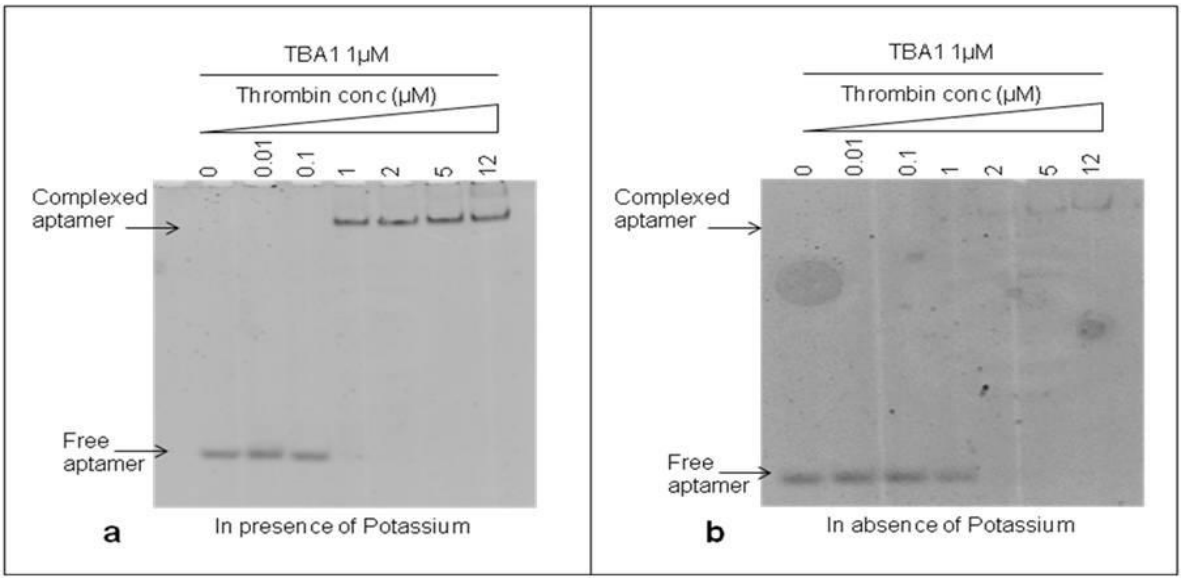


Figure 2.19 Electrophoretic Mobility Shift Assay (EMSA) of unmodified TBA1 and thrombin in presence (a) and in absence of Potassium (b). TBA1 was incubated with the protein under the conditions described. We have indicated in the figure the respective TBA1 and thrombin concentrations. A control reaction without thrombin was performed in all experiments. KCl was added to the polyacrilamide gel and to the running buffer. Binding reactions were applied on a 12% non-denaturing PAA gel containing 1X TBE buffer and KCl 10 mM.

The mobility of free and complexed aptamers, stained by SybrGreen II®, was detected using the Geliance 600 Imaging System.

We can observe in Figure 2.19a that the band corresponding to free TBA1 shifts up in the gel starting from 1:1 TBA1/thrombin ratio, consistent with the lower mobility of the TBA1-thrombin complex. At the same time, the free TBA1 band disappears. The presence of the upper band with lower electrophoretical mobility indicates that TBA1 is complexed with the target protein under these conditions, *i.e.*, in the presence of K^+ ions. Figure 2.19b shows the result of the same experiment performed in the absence of K^+ ions at all steps (folding, incubation with thrombin and analysis by EMSA). We can observe that TBA1 binds thrombin but with reduced affinity, since the shift of TBA1 band appears at a higher TBA1/thrombin ratio (1:2). Although thrombin itself is known to act as a molecular chaperone promoting the G-quadruplex formation in the aptamer [117], the presence of K^+ cations favours G-4 formation and facilitates TBA1-thrombin binding. Higher concentrations of potassium ions were also analyzed (up to 100 mM), but the high salt in the gel and in the running buffer impairs EMSA assay. For these reasons we considered as optimal binding conditions the presence of K^+ ions and 10 mM KCl was included in our systems.

Role of TBA1 Modifications

To immobilize TBA1, the capture layer in our aptamer-based microarray, 5'-amino modification of the oligonucleotide was required for aptamer anchorage on the glass slides. Therefore, we investigated in solution the effect of this modification on thrombin binding: folded TBA1-NH_2 (1 μM) was incubated with increasing concentrations (0, 0.001, 0.01, 0.1, 1, 2, 5 and 12 μM) of human thrombin following the procedure described above for unmodified TBA1.

Figure 2.20a shows that, starting from a 1:1 aptamer to thrombin ratio, the band of the free aptamer gradually disappears and at the same time the band with lower mobility corresponding to TBA1-NH_2-thrombin complex appears. Hence, comparing these results with what seen for unmodified TBA1 (Figure 2.19a), the 5'-amino modification does not impair the TBA1-thrombin binding.

To evaluate the effect of the polyT spacer added at the 5'-end of the oligonucleotide, the same experiment was performed using TBA1(12T)NH_2: the results are reported in Figure 2.20b. From a 1:1 aptamer/thrombin ratio the band of the free aptamer shifts up in the gel in bands that are slightly smeared compared to those of unmodified aptamer-thrombin (compare Figure 2.20 with Figure 2.19a). Although more than one complex seems to be present relative

to the unmodified control, suggesting the possible presence of different folding equilibria, the addition of the polyT (12T) spacer at the 5'-end of the 15 nucleotides capture aptamer TBA1 does not inhibit the recognition of the protein in solution, in agreement with what reported in solid phase for the recognition of Cy-labeled thrombin with the modified aptamer [121].

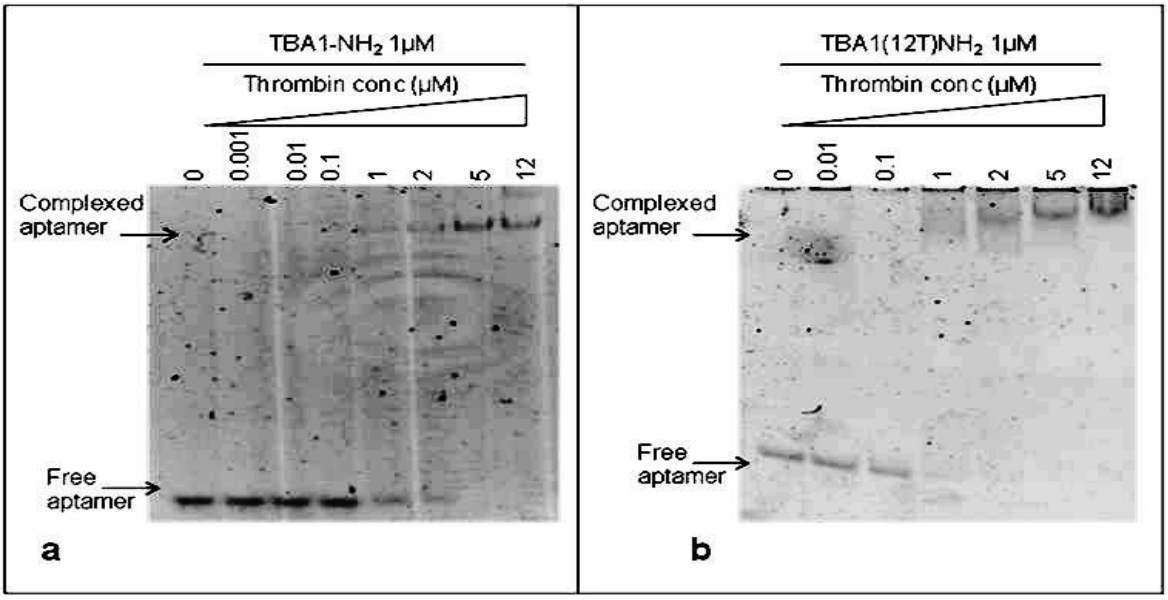


Figure 2.20 Electrophoretic Mobility Shift Assay (EMSA) of TBA1-NH_2 with thrombin (a) and of TBA1(12T)NH_2 with thrombin (b). TBA1-NH_2 and TBA(12T)NH_2 were separately incubated with the protein under the conditions described. We have indicated in the figure the respective aptamer and thrombin concentrations. A control reaction without thrombin
was performed in all experiments. Binding reactions were applied on a 12% non-denaturing PAA gel containing 1X TBE buffer and KCl 10mM. The mobility of free and complexed aptamers, stained by SybrGreen II®, was detected using the Geliance 600 Imaging System.

2.4.1.1.2 TBA2-Thrombin Interaction in Solution

Evaluation of Labeling with Alexa Fluor 488 or Cy5 Dye

The 29-mer DNA thrombin binding aptamer TBA2, planned to be used for fluorescence detection in our aptamer-based microarray, has to bear important chemical modifications at its 5' end: to rule out any possible interference of the aromatic systems of fluorescent dyes on the folding and on the recognition of the aptamer with the target, thrombin binding by the unmodified TBA2 was compared to that exhibited by TBA2 fluorescently labeled at the 5'-end with the green light emitter Alexa Fluor 488 (TBA2-Alexa) and the red light emitter Cy5 (TBA2Cy5). The gel shifts of the unmodified and modified aptamers are shown in Figure 2.21.

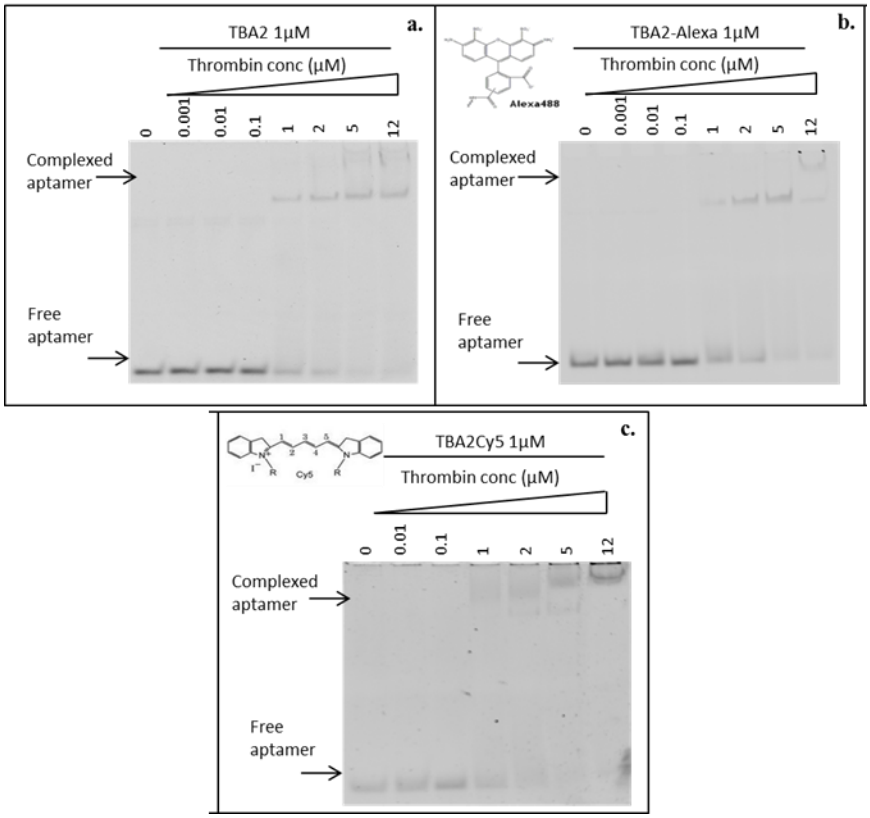


Figure 2.21 Electrophoretic Mobility Shift Assay (EMSA) of unmodified TBA2 with thrombin (a), TBA2-Alexa with thrombin (b) and TBA2Cy5 with thrombin (c). Each aptamer was incubated with the protein under the conditions described. We have indicated in the figure the respective aptamer and thrombin concentrations. A control reaction without thrombin was performed in all experiments. Binding reactions were applied on a 12% non-denaturing PAA gel containing 1X TBE buffer and KCl 10 mM. The mobility of free and complexed aptamers, stained by SybrGreen II® was detected using the Geliance 600 Imaging System.

The unmodified TBA2 in KCl 10 mM folds in G-quadruplex, as reported by CD titrations [148], and binds thrombin from TBA2/thrombin ratio 1:1 (Figure 2.21a): the band corresponding to TBA2 shifts up and new bands with lower mobility, representing the TBA2-thrombin complex, appear in the upper part of the gel, as expected.

The same result was obtained analyzing the 5'-Alexa Fluor 488 labeled TBA2 interaction with thrombin: we can observe in Figure 2.21b that TBA2-Alexa shifts up at the same aptamer/protein ratio of the unlabeled aptamer used as control (Figure 2.21a) and at the highest thrombin concentration a complex with even lower mobility appears. The 5'-labeling of TBA2 with Alexa Fluor 488 does not interfere with the binding affinity, but at high protein concentrations the planar surfaces individuated by the fluorophore could determine the formation of complexes with a different stoichiometry or a different shape.

An evident gel shift at same aptamer/protein ratio compared to the control was obtained even when the 5'-Cy5 labeled TBA2 was analyzed (Figure 2.21c). The 5'-labeling of TBA2 with Cy5 evidently influences the overall hydrodynamic volume of the complex with

thrombin, as seen by the appearance of bands of relative lower mobilities (Figure 2.21c) compared to the complex with unlabeled aptamer controls (Figure 2.21a). The planar and hydrophobic surfaces individuated by the fluorophore and, possibly, also the structure(s) of the binary complex(es) with thrombin are affected: at the highest protein/oligo ratio the modified aptamer forms complexes with even lower electrophoretic mobilities, possibly consistent with the formation of aptamer/protein complexes with a different stoichiometry or a different shape. These observations are in line with literature data on the strict structural requirements of TBA2 for correct folding and target binding [149]. However, we can observe that the presence of the labeling dye Alexa Fluor 488 and Cy5 conjugated to the aptamer allows thrombin recognition, since the binary complex formation is not abolished. The solution experiments point to the requirement of careful handling of the detection aptamer when adapting the solution studies to the solid phase: correct structuring of the labeled TBA2 must be ensured by denaturing/folding protocols prior to its use to ensure thrombin recognition.

2.4.1.1.3 Supershift Assay: Optimization of TBA1-Thr-TBA2 Interaction

After having completed the studies of thrombin binding for each set of modified aptamers, we verified the formation of the ternary complex in solution. The Supershift Assay using the aptamers with the modifications needed for solid phase will indicate the feasibility of the sandwich system in the planned microarray and will constitute the proof of principle of our approach.

The Supershift Assay was performed with two combinations of aptamers: TBA1-NH_2 as capture aptamer and TBA2-Alexa as detection aptamer, using human thrombin as protein target; TBA1(12T)NH_2 as capture layer and TBA2Cy5 as detection layer, using human thrombin and Alexa555-labeled human thrombin to verify ternary complex formation. The primary aptamer and the secondary aptamer were folded in KCl, mixed and then incubated with thrombin(s) in the Binding Buffer at 25°C for 30 min. After incubation, free aptamer and thrombin-aptamer complexes were resolved by 12% non-denaturing polyacrylamide gel containing KCl. Free aptamers and the binary complex of each aptamer with thrombin were loaded in the gel as controls.

TBA1-NH_2-Thr-TBA2-Alexa ternary complex

The formation of the TBA1-NH_2-Thr-TBA2-Alexa ternary complex can be clearly visualized in Figure 2.22: a new band with different mobility from the binary complexes appears, while the binary complexes bands disappear from their position.

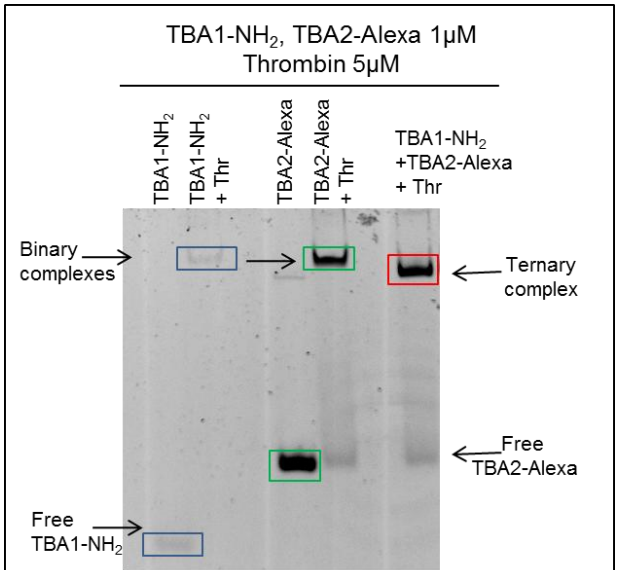


Figure 2.22 Electrophoretic Mobility Supershift Assay (EMSA) of TBA1NH_2 + thrombin + TBA2-Alexa. Each aptamer was incubated separately or simultaneously with thrombin under the conditions described. We have indicated in the figure the respective aptamer and thrombin concentrations. A reaction control without thrombin was performed for each aptamer. Binding reactions were applied on a 12% non-denaturing PAA gel containing 1X TBE buffer and KCl 10 mM. The mobility of free and complexed aptamers, stained by SybrGreen II® was detected using the Geliance 600 Imaging System.

Free aptamers and the binary complex of each aptamer with thrombin were loaded in the gel as controls. When TBA1-NH_2 and TBA2-Alexa were incubated simultaneously with thrombin a new band with higher mobility in the native polyacrylamide gel system appears. The molecular weight of the ternary complex is sure higher than that of the binary complexes, but the ternary complex, in this case, runs into the gel faster than the binary ones. It is possible that the ternary complex has a hydrodynamic volume which favors the penetration into the gel: the TBA1-NH_2-Thr-TBA2-Alexa sandwich probably owns, in solution, a compact shape and a negatively charged surface, due to the presence of the Alexa488 fluorophore that is double negatively charged. Hence, we can safely presume that the new band corresponds to the sandwich complex TBA1-NH_2-Thr-TBA2-Alexa.

TBA1(12T)NH$_2$-Thr-TBA2Cy5 ternary complex

The Supershift Assay was performed with another pair of aptamers: TBA1(12T)NH$_2$ as capture layer and TBA2Cy5 as detection layer; human thrombin and Alexa555-labeled human thrombin were used to verify ternary complex formation.

The results of this experiment employing different detection systems are reported in Figure 2.23.

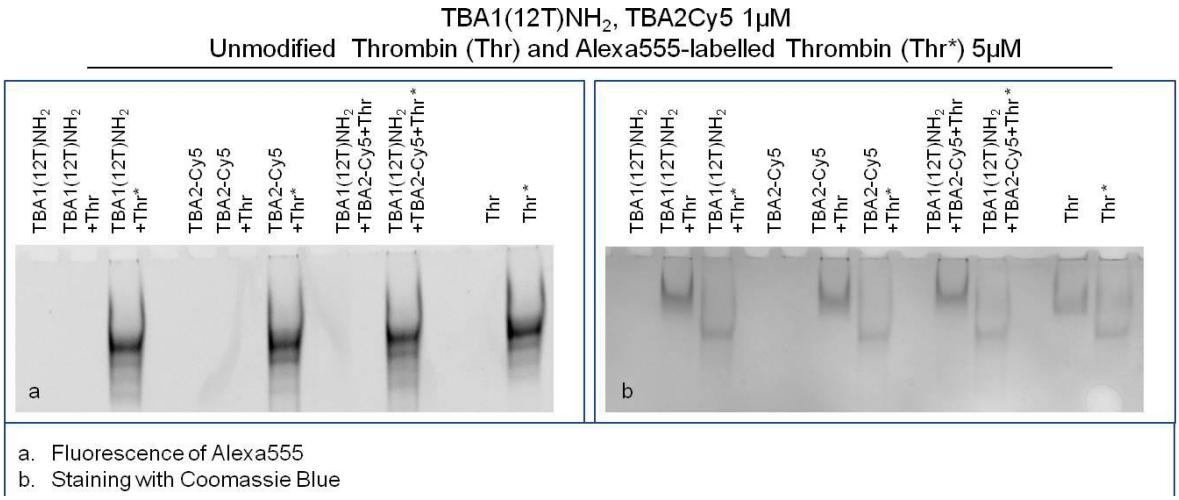


Figure 2.23 Electrophoretic Mobility Supershift Assay (EMSA) of TBA1(12T)NH2 + thrombin + TBA2Cy5: detection of Alexa555-labeled thrombin (a); detection of thrombin with Coomassie Blue staining (b). Each aptamer was incubated separately or simultaneously with the unmodified and fluorescently labeled protein under the conditions described. We have indicated in the figure the respective aptamer and thrombin concentrations. A reaction control without thrombin was performed for each aptamer; other reaction controls were the unmodified and fluorescently labelled thrombin incubated without any aptamer. Binding reactions were applied on a 12% non-denaturing PAA gel containing 1X TBE buffer and KCl 10 mM. The mobility of free and complexed aptamers was detected using the Geliance 600 Imaging System.

Figure 2.23a shows the image detected analyzing green light emission without any gel staining: the bands detected correspond to thrombin labelled with Alexa555. Figure 2.23b shows the result of the supershift assay using Coomassie Blue to visualize the position of unlabeled thrombin and Alexa555 labeled thrombin, either alone or complexed with aptamers. A third method of staining with the fluorescent intercalator SybrGreen II (not shown), allows the detection of the aptamer bands in the gel.

Thrombin has an almost neutral pI [147] and its mobility is limited in the native gel system employed (Tris borate buffer, pH 8.3). It is clearly evident in Figure 2.23b that Alexa labeling changes the mobility of the protein: Alexa-thrombin moves faster compared to unmodified thrombin. Although the molecular structure of Alexa555® is not disclosed by the manufacturer, it possesses a negative charge (information provided by Invitrogen) that alters

the charge profile of the protein, and is responsible for the higher mobility of the protein in the native polyacrylamide gel system used. The band of each free aptamer shifts up in presence of fluorescently labeled or unmodified thrombin, due to the higher molecular weight aptamer-protein complex formation, as verified in the previous experiments. When in complex with both aptamers thrombin forms a ternary complex verified by the appearance of the supershift band with lower mobility than the binary complex band (Figure 2.23b). This is clearly evident in the case of unmodified thrombin, and it is also verified for Alexa-thrombin. The relative mobilities of the complexes, calculated as Rf for each band, are summarized in Table 2.4.

	+ TBA1 (12T)NH_2		+ TBA2Cy5		+ TBA1(12T)NH_2 + TBA2Cy5			
Protein	Thr	Alexa555-Thr	Thr	Alexa555-Thr	Thr	Alexa555-Thr	Thr	Alexa555-Thr
Rf	0.05	0.09	0.05	0.09	0.04	0.08	0.07	0.09

Table 2.4 Relative mobilities (Rf) of binary and ternary complexes.

Alexa-labeled protein moves faster than the unlabeled thrombin (Rf: 0.09 *vs.* 0.07); in the case of binary complexes with human thrombin the bands are shifted up from the protein alone (to 0.05 for both capture and detection aptamer), and are further shifted (to 0.04) when all the system players are in a ternary complex. With Alexa-thrombin the mobilities of the binary complexes approximate that of the labeled protein (Rf = 0.09). However, when both aptamers are incubated with Alexa-thrombin the ternary complex band mobility is retarded to an Rf of 0.08. Therefore, in the context of the system corresponding to the solid phase sandwich the solution study indicates that the ternary complexes are formed in all cases investigated and the system can be applied to the solid phase.

2.4.1.2 Analysis in Solid Phase

2.4.1.2.1 Optimization of the Aptamer-Thrombin Sandwich Protocol

The sandwich system with the modified aptamers analyzed in solution was then applied to the microarray system, *i.e.*, in solid phase.

The SAM was performed first with the aptamer system shown in Figure 2.22, employing TBA1-NH_2 as capture layer and TBA2-Alexa as detection layer. In this case the sandwich formation is detected by the green fluorescence of Alexa Fluor 488 conjugated to the oligonucleotide TBA2.

Alexa647-labelled thrombin was also used as control to further verify the sandwich formation by the yellow fluorescence resulting from the co-localization of green (Alexa488 conjugated to the aptamer) and red (Alexa647 conjugated to the protein) fluorophores.

The protocol of the SAM consisted, in this case, in deposing on the capture layer TBA1-NH_2 thrombin (unmodified or Alexa647-labeled thrombin) pre-incubated with the detection aptamer. The samples (six spots for each sample) were prepared as shown schematically in Table 2.5; each sub-array was physically isolated from the others.

Sub-array	1	2	3	4
Capture layer	TBA1-NH_2 OTA	TBA1-NH_2 OTA	TBA1-NH_2 OTA	TBA1-NH_2 OTA
Protein	Thr	Thr	Alexa647-Thr	Alexa647-Thr
Detection layer	TBA2-Alexa	-	TBA2-Alexa	-

Table 2.5 Schematic representation of the samples incubated in the Sandwich Aptamer Microarray.

The results of the SAM experiments is shown in Figure 2.24.

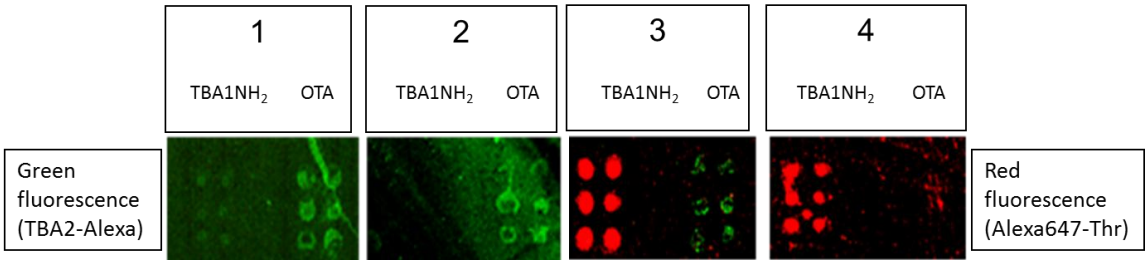


Figure 2.24 Images of the slide after the Sandwich Aptamer Microarray (SAM). In each subarray $TBA1NH_2$ was printed on the left, while OTA on the right. The fluorophore Alexa488 emitted green light, while Alexa647 emitted red light. For this reason red fluorescence represents labeled thrombin and green fluorescence represents TBA2-Alexa.

In chambers 1 and 2 the unmodified thrombin was used. The TBA1-NH_2-Thr-TBA2-Alexa sandwich formation would be detected in chamber 1 by the green fluorescence of Alexa Fluor 488 conjugated to TBA2, but no significant fluorescence signal is noted. The fluorescence detected in chamber 2 in which thrombin was deposed on the capture layer without the detection aptamer, represents the background signal: no differences are observed comparing chamber 1 and chamber 2, and therefore it is clear that the sandwich formation is not detected.

Alexa-labeled thrombin allows a further control for thrombin presence in the sandwich. In chamber 3 the sandwich signal would be represented by the yellow fluorescence resulting from the co-localization of Alexa-labeled thrombin and TBA2-Alexa, but the fluorescence signal detected on TBA1-NH_2 is red, corresponding to the labeled thrombin and is the same fluorescence detected in the chamber 4 in which labeled thrombin was incubated without the detection aptamer. Since no thrombin signal is detected on the negative control OTA, the ability of TBA1-NH_2 to capture thrombin was confirmed, but it is not enough to attend the sandwich formation. The recognition of TBA2-Alexa with human thrombin seems to be the critical step in the development of the SAM system in solid phase.

Hence, the SAM protocol using TBA1-NH_2 as capture layer and TBA2-Alexa as detection layer results not successful for thrombin detection and this pair of aptamers was abandoned.

The SAM was then performed with the aptamer system shown in Figure 2.23, *i.e.*, using TBA1(12T)NH_2 as capture layer for thrombin and the fluorescently labeled TBA2Cy5 as detection layer. In this case the sandwich formation is detected by the red fluorescence of TBA2Cy5.

Alexa555-labelled thrombin was also used as control to further verify the sandwich formation by the yellow fluorescence resulting from the co-localization of green (Alexa conjugated to the protein) and red (Cy conjugated to the oligonucleotide) fluorophores. Two different protocols were tested for the sandwich formation on surface: a one-step procedure by deposing on the capture layer thrombin(s) pre-incubated with the detection aptamer (chambers 1 and 2) and a two-steps procedure deposing thrombin on the capture layer, followed by a final deposition of the detection layer. The slides carefully rinsed with the PBS Buffer and MilliQ H_2O were finally spin-dried and scanned. The samples (six spots for each sample) were prepared as shown schematically in Table 2.6; each sub-array was physically isolated from the others.

Sub-array	1	2	3	4
Capture layer	TBA1 (12T)NH_2 OTA	TBA1 (12T)NH_2 OTA	TBA1 (12T)NH_2 OTA	TBA1 (12T)NH_2 OTA
Protein	Thr (0.5µM) pre-complexed with TBA2Cy5(1µM)	Alexa555-Thr (0.5µM) pre-complexed with TBA2Cy5(1µM)	Thr (0.5µM)	Alexa555-Thr (0.5µM)
Detection layer			TBA2Cy5 (1µM)	TBA2Cy5 (1µM)
Procedure	One-step		Two-steps	

Table 2.6 Schematic representation of the samples incubated in the SAM protocol.

The results of the SAM experiments are shown in Figure 2.25.

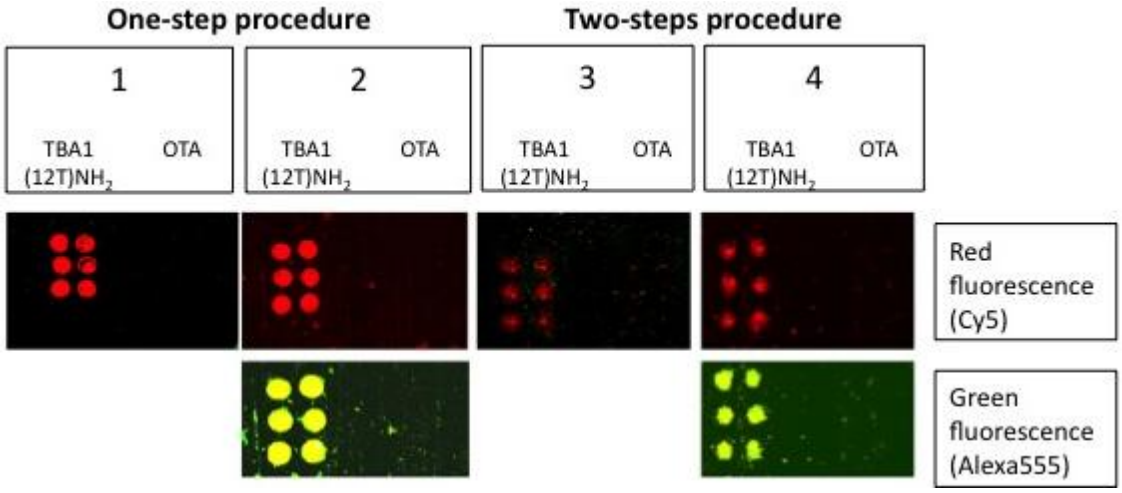


Figure 2.25 Images of the slide after the Sandwich Aptamer Microarray (SAM). In each subarray TBA1(12T)NH2 was printed on the left, while OTA on the right. The fluorophore Cy5 emitted red light, while Alexa555 emitted green light. For this reason red fluorescence represents TBA2Cy5 and green fluorescence represents labeled thrombin. In chambers 2 and 4, in which labeled thrombin was incubated, the fluorescence was detected with two different contrasts to reveal first red light and then green light: yellow spots indicate the simultaneously localization of the two fluorophores.

In chambers 1 and 2 the red spots (TBA2Cy5) in correspondence of the selection aptamer TBA1(12T)NH_2 anchored on the glass slide reveal that the sandwich was formed with high specificity for the thrombin aptamer: the detection of the Cy5 labeled aptamer is reproducible across the six spots of each sub-array. The recognition is also specific: no signal is detected for the negative control OTA. SAM is formed with thrombin and also with Alexa-labeled thrombin, consistently with the solution studies: in the sub-arrays 2, Alexa-labeled thrombin allows a further control for thrombin presence in the sandwich by the yellow fluorescence

resulting from the simultaneous presence of green fluorescence (Alexa555-thrombin) and red fluorescence (TBA2Cy5). The absence of unexpected hybridization between TBA1 and TBA2Cy5 and the formation of a specific complex between Alexa-thrombin and immobilized TBA1 were also verified independently (not shown).

The one-step protocol is more efficient than the two step procedure: the detection of the sandwich is sharp and specific in chambers 1 and 2 compared with chambers 3 and 4. In chamber 3 the signal obtained is weak, while chamber 4 (Alexa-labeled protein) exhibits a higher background and lower selectivity at this ionic strength. As indicated by the solution studies, the recognition of TBACy5 with human thrombin seems to be the critical step in the development of the best protocol for the solid phase system. The one-step protocols ensures enough time to the modified detection aptamer for the formation of the correct binary complex able to expose the protein epitopes recognized by the immobilized capture aptamer. The presence of the fluorescent dye conjugated to the protein has also a negative effect, possibly influencing the pattern mode of recognition aptamer-thrombin, reflected in the loss of specificity of binding. The microarray system protocol of chamber 1 results a successful application of the aptamer sandwich devised.

2.4.1.2.2 Optimization and characterization of the SAM for thrombin detection

The sandwich formation in solid phase was verified exploiting the two non-overlapping DNA aptamers TBA1(12T)NH_2 and TBACy5, recognizing different exosites of human thrombin [145]. Moreover, the correct sandwich formation was confirmed by the fluorescence signals of the sandwich and the recognition by aptamers was specific (Figure 2.25).

The SAM system for thrombin detection was further optimized and characterized as described below, to be developed as a simple and sensitive thrombin diagnostic tool.

Thrombin specificity

With the aim of optimizing our system in view of future applications in multiplex microarray systems, we enlarge our analysis of the aptasensor specificity incubating the TBA1(12T)NH_2 microarray with proteins related and unrelated to human thrombin. Specificity toward unrelated proteins was evaluated using BSA and lysozyme [123], while human VEGF165 was chosen for system specificity evaluation since it has both a fibrinogen [150] and a heparin binding domain (FBD and HBD respectively) like thrombin [151].

Figure 2.26 shows that no significant fluorescent signal was recorded on TBA1(12T)NH_2 microarray spots after incubation of samples containing BSA or lysozyme pre-incubated with TBA2Cy5, confirming the specificity of the two aptamers towards human thrombin. The aptasensor is able to discriminate between related target domains: experimental data in Figure 2.26 shows clearly that no significant binding on TBA1(12T)NH_2 spots was observed in presence of VEGF 165.

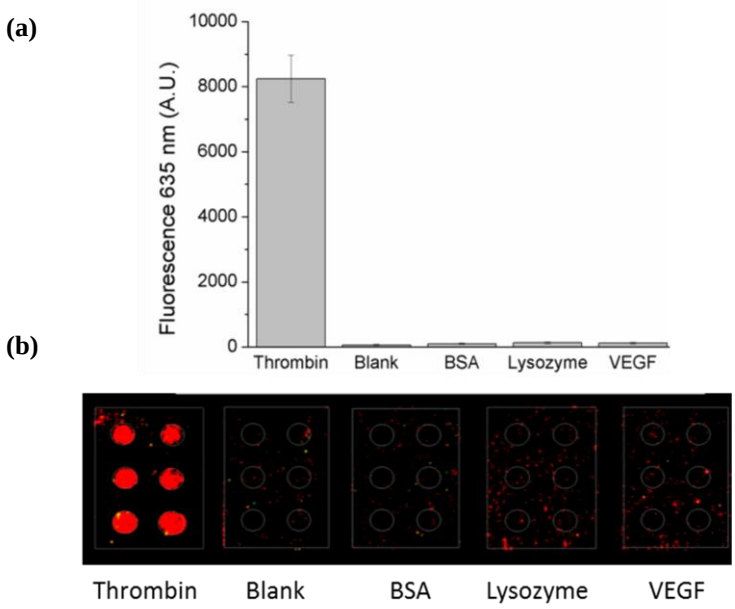


Figure 2.26 Specificity of aptasensor for thrombin. Fluorescence signal (635 nm) recorded on thrombin aptamer microarray after the incubation of the different indicated proteins with TBA2Cy5 (protein and TBA2 final concentration: 500 nM and 1 mM, respectively); **(a)** thrombin, blank (TBA2Cy5 only), BSA, lysozyme and VEGF and **(b)** corresponding microarray images.

These results hence confirm that TBA1 and TBA2 used in the two-site binding assay ensure high recognition specificity and are able to discriminate between HBD and FBD from different proteins.

Effect of incubation time

Assay incubation time was optimized in order to obtain the maximum fluorescent signal from analyzed samples and at the same time to perform the assay within a time compatible with routine analysis. After a pre-incubation step between the protein and the detection aptamer TBA2Cy5, samples were incubated on the TBA1(12T)NH_2-printed microarray slide for a time ranging from 5 minutes up to 180 minutes. As shown in Figure 2.27 the fluorescent signal from TBA2Cy5 increases exponentially over time, reaching a maximum intensity after 180 minutes, but the difference between the relative fluorescence intensity at 2 hours and 3 hours incubation time is lower than 6%. To achieve high assay sensitivity and to perform the

assay in an overall time compatible with that of standard commercial assays, we performed all the following experiments using a 2 hours incubation time.

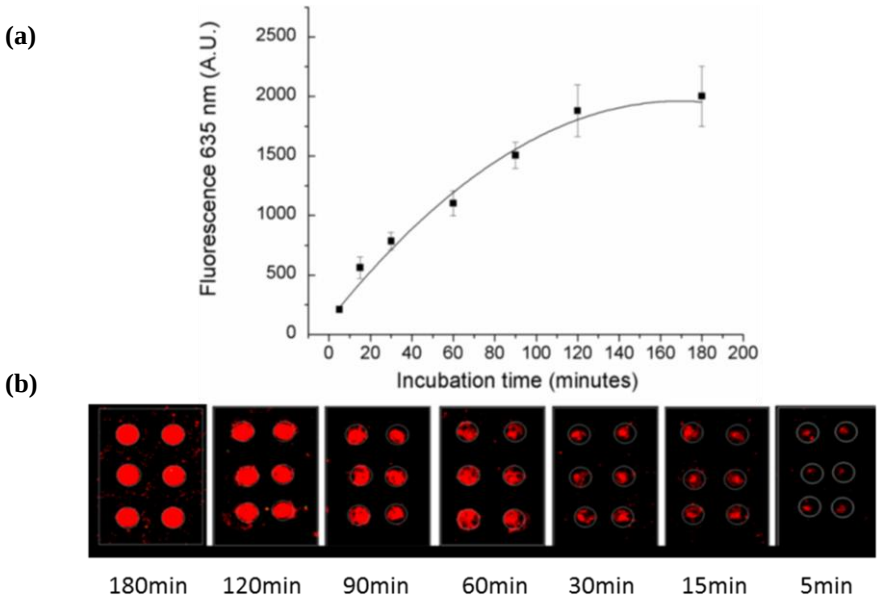
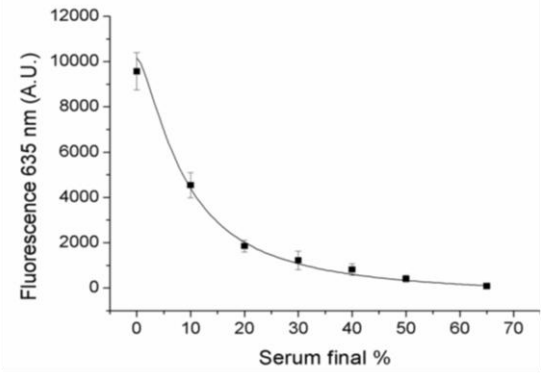


Figure 2.27 Effect of incubation time. Fluorescence signal (635 nm) recorded on TBA1(12T)NH_2 microarray **(a)** after the incubation of thrombin and TBA2Cy5 for different times (180-120-90-60-30-15-5 minutes) and **(b)** corresponding microarray images.

Effect of serum

Immunoassays protocols require sample dilutions prior to analysis in order to dilute possible interferents [152]. To evaluate the system sensing ability in complex matrixes, equal thrombin amounts (500nM) were analyzed in the presence of fetal bovine serum (FBS) at different dilutions ranging from 0% to 65%. After a short pre-incubation in the presence of FBS, samples were incubated for two hours on the array slide printed with TBA1(12T)NH_2. Collected data shown in Figure 2.28 demonstrate that the system can accurately work in the presence of 20% as well as 10% FBS without affecting the system performance. A 20% FBS solution, corresponding to a biological sample dilution of 1:5 is therefore compatible with the sandwich aptamer microarray analysis.

(a)



(b)

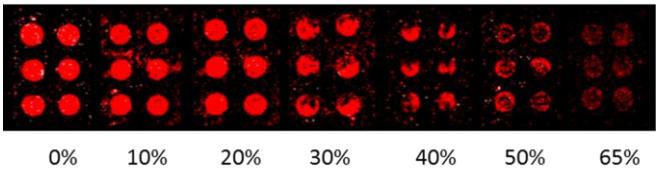


Figure 2.28 Effect of serum. Fluorescence signal (635 nm) recorded on TBA1(12T)NH_2 microarray **(a)** after the incubation of thrombin and TBA2Cy5 complexes in presence of different FBS dilutions (0%-10%-20%-30%-40%-50%-65%) and **(b)** corresponding microarray images.

Limit of detection of direct and indirect methods

The microarray developed and described is based on the "direct" detection of the labeled secondary aptamer recognizing the analyte-aptamer complex on the slide. This system was compared to an "indirect" method based on the formation of a similar sandwich-type on the microarray, but employing a secondary aptamer conjugated with biotin, followed by a detection step with fluorescently labeled streptavidin. This method would allow the development of a multiplex systems employing biotinylated aptamers specific for each analyte followed by a single detection step with labelled streptavidin. We set up and compared both methods for the thrombin microarray, and evaluated the assay limit of detection (LOD) and limit of quantification (LOQ) in both cases: different thrombin concentrations (200 nM - 100 nM - 50 nM - 10 nM - 5 nM – 1 nM - 0 nM) were pre-incubated with the two different secondary aptamers (respectively TBA2Cy5 and TBA2-biotin) and finally laid on the TBA1(12T)NH_2 microarray slide. In the first (direct) case, the microarray is analysed at 635 nm fluorescent excitation to detect Cy5 directly conjugated to the aptamer, in the second ("indirect") method the microarray is analysed at 532 nm fluorescent excitation to detect Cy3-labelled streptavidin laid over the treated samples. For each dilution, the mean (subtracted to background values) of six spots was calculated, with corresponding standard deviation.

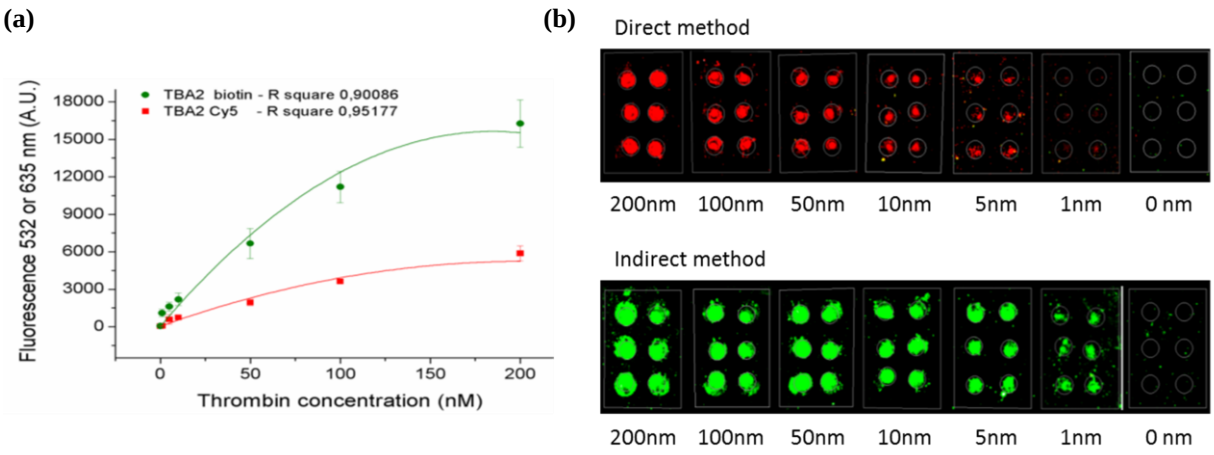


Figure 2.29 Direct and indirect method for thrombin detection. **(a)** Fluorescence signal (635 or 532 nm) plot of thrombin aptamer microarray after the incubation of thrombin and TBA2Cy5 (direct method, red squares, 635 nm) or thrombin and TBA2-biotin plus Cy3-streptavidin (indirect method, green dots, 532 nm) and **(b)** relative microarray images. Tested thrombin dilutions were: 200 nM – 100 nM - 50 nM - 10 nM - 5 nM - 1 nM - 0 nM, TBA2Cy5 (red squares) or TBA2-biotin (green dots) concentration was 400 nM. For the indirect method the microarray slide was incubated for 1 hour with Cy3-streptavidin, for a total incubation time of two hours in all cases.

As shown by the red squares symbols in Figure 2.29a, the direct method (red dots) employing TBA2Cy5 yields a LOD of 0.85 nM and LOQ of 2 nM, corresponding to 31.45 ng/mL and 74 ng/mL of thrombin, respectively. In the indirect method with TBA2-biotin detected by fluorescent streptavidin (green dots), a LOD of 0.55 nM and LOQ of 1.35 nM were obtained, corresponding to 20.35 ng/mL and 49.95 ng/mL of thrombin, respectively.

Indirect method for aptasensor detection: effect of serum

To complete our characterization of the aptamer microarray for thrombin detection based on the indirect method of fluorescent analysis, samples with decreasing concentrations of thrombin were analysed in solutions containing serum, in order to evaluate the sensing ability to detect thrombin in a complex matrix. Samples with decreasing concentrations of thrombin (50 nM - 10 nM - 5 nM - 1 nM - 0.5 nM - 0 nM) and 20% FBS, were pre-incubated in presence of 200 nM of TBA2-biotin and finally incubated on the microarray slide as previously described. After Cy3-labelled streptavidin incubation, data were collected and reported in Figure 2.30.

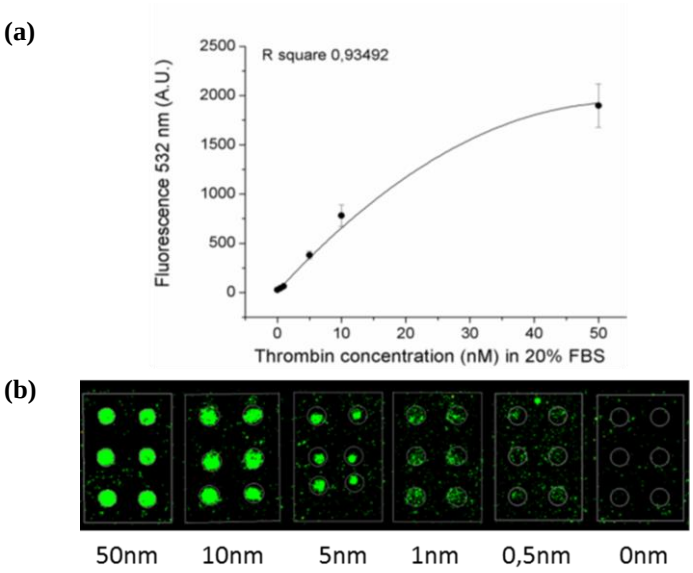


Figure 2.30 Thrombin calibration curves in FBS solution: **(a)** fluorescence signal (532 nm) plot of TBA1 microarray in 20% FBS solution after the incubation of thrombin and TBA2-biotin plus Cy3-streptavidin and **(b)** corresponding microarray images. Tested thrombin dilutions were: 50 nM - 10 nM - 5 nM - 1 nM - 0.5 nM - 0 nM, with a fixed TBA-2 biotin concentration of 200 nM.

The presence of 20% serum contributes to lower the overall specific detection signal, as expected, but the LOD and LOQ, respectively 0.53 nM and 1.25 nM, allows sub-nanomolar detection of the analyte also in these conditions. Data collected from thrombin calibration curve in presence of 20% FBS demonstrate that the developed system can accurately work also in the presence of a complex matrix, ensuring LOD and LOQ overlapping to those one obtained in the presence of pure protein samples.

Hence, two transducing and recognition strategies for thrombin detection were optimized. With the aim of improving our system in view of future applications in multiplex microarray systems, we enlarge our analysis on other target proteins.

2.4.2 Human $VEGF_{165}$ detection through a Sandwich Aptamer Microarray

2.4.2.1 Analysis in Solution

2.4.2.1.1 Vap7-$VEGF_{165}$ Interaction in Solution

Vap7 is a recently selected DNA aptamer against $VEGF_{165}$ and its target epitope is the receptor-binding domain (RBD) of $VEGF_{165}$ (K_d value of 20 nM) [141]. Vap7 is the first VEGF binding aptamer that folds into G-quadruplex structure: the K^+ cation has been reported to stabilize the G-quadruplex structure of nucleic acids [153] and in particular K^+ is important for the structure of Vap7 when it binds to VEGF [141]. For this reason, to favor and conserve the Vap7 G-quadruplex structure, we decided to perform the analysis of Vap7-VEGF interaction in the presence of KCl at all steps (folding, incubation with VEGF and analysis by EMSA).

We first investigated the binding of unmodified Vap7 aptamer to $VEGF_{165}$.

The binding of Vap7 to $VEGF_{165}$ is not well characterized in literature; therefore, we investigated the suitable conditions to observe the aptamer-protein binding by EMSA. In the case of $VEGF_{165}$ is very important the choice of the buffer in the gels: a buffer with a pH>pI of the protein is needed to have the protein negatively charged and so that the protein can enter into the gel systems. $VEGF_{165}$ has a pI of 8.5 and therefore a strongly basic buffer is required: glycine 50 mM pH 9.5 added of KCl was used as buffer in the gels and as running buffer for EMSA.

Vap7 was denatured, folded and incubated with the protein (up to 10 µM) in TBSE containing additional potassium and the complex formation was analyzed by EMSA including potassium in the buffer system. Two different techniques were used to detect free and complexed aptamer in the gel: in Figure 2.31a is reported the result obtained by staining with SybrGreen II; in figure 2.31b is reported the result obtained by using the [γ^{32}P] ATP radiolabelled Vap7 (Vap7*) aptamer, in order to acquire more accurate signals of the aptamer in the gel.

In both cases, the band corresponding to free aptamer shifts up in the gel starting from 1:5 Vap7/VEGF ratio and gradually disappears; at the same time the band with lower mobility corresponding to the aptamer-protein binary complex appears. Although the bands in the gel are evidently smeared, probably due to the high salts concentration that often determines this effect in the electrophoresis assays, the recognition of $VEGF_{165}$ by Vap7 is verified.

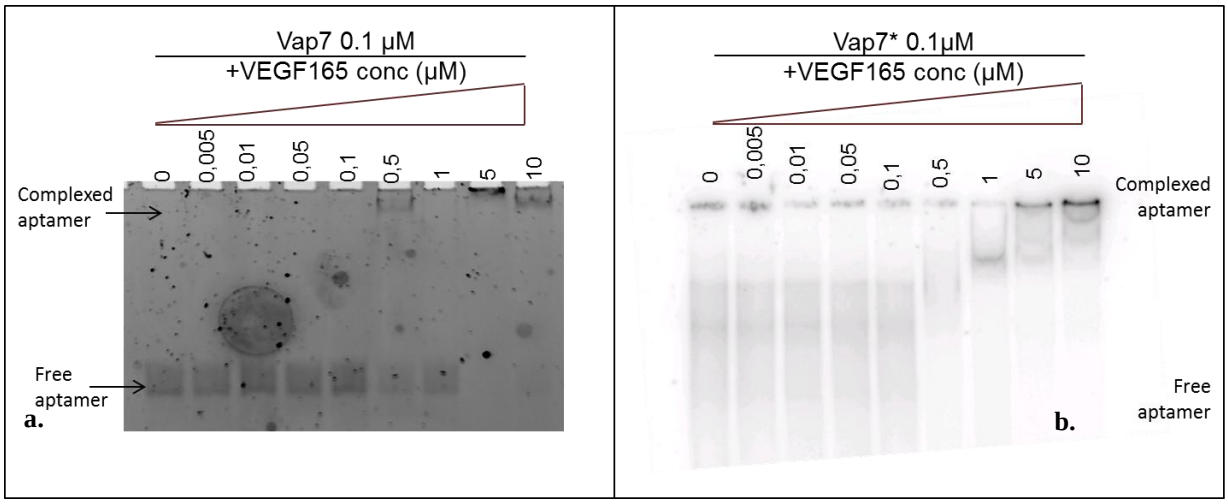


Figure 2.31 Electrophoretic Mobility Shift Assay (EMSA) of unmodified Vap7 with VEGF165 (a) and of [γ^{32}P]ATP-labeled Vap7 with VEGF$_{165}$ (b). Vap7 and Vap7* were separately incubated with the protein under the conditions described. We have indicated in the figure the respective aptamer and VEGF concentrations. A control reaction without VEGF was performed in all experiments. Binding reactions were applied on a 12% non-denaturing PAA gel containing 1X glycine buffer and KCl 10mM. The mobility of free and complexed aptamer, stained by SybrGreen II®(a), was detected using the Geliance 600 Imaging System. The mobility of free and complexed aptamer, radioactively labeled (b), was detected using the *Phosphorimager* STORM B40.

No information was known about Vap7 employment as capture layer or detection layer in sensing system, *i.e.* in solid phase. Therefore, prior to the development of a Sandwich Aptamer Microarray (SAM) for VEGF$_{165}$ detection, we investigated the effect of both the 5'-amino and the 5'-FAM modifications on protein binding. Folded Vap7-NH$_2$ or Vap7-FAM was incubated with increasing concentrations of VEGF$_{165}$ following the procedure described above for unmodified Vap7 (Figure 2.32).

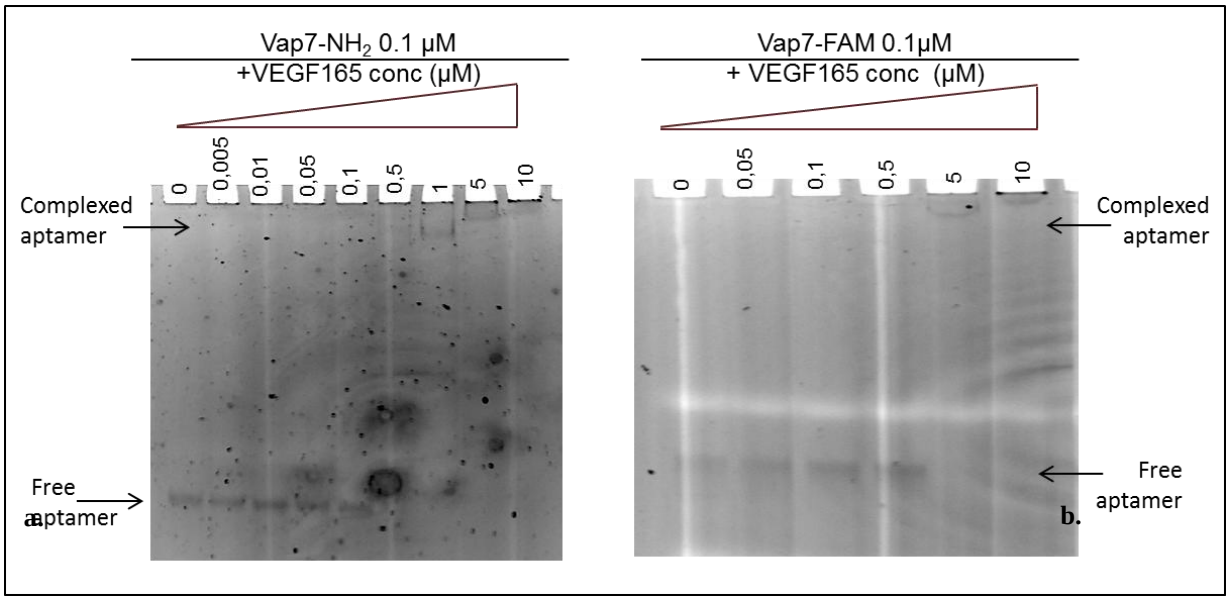


Figure 2.32 Electrophoretic Mobility Shift Assay (EMSA) of Vap7-NH$_2$ with VEGF$_{165}$ (a) and of Vap7-FAM with VEGF$_{165}$ (b). Vap7-NH$_2$ and Vap7-FAM were separately incubated with the protein under the

conditions described. We have indicated in the figure the respective aptamer and VEGF concentrations. A control reaction without VEGF was performed in all experiments. Binding reactions were applied on a 12% non-denaturing PAA gel containing 1X glycine buffer and KCl 10mM. The mobility of free and complexed aptamer, stained by SybrGreen II®(a), was detected using the Geliance 600 Imaging System.

Figure 2.32a shows that, starting from a 1:5 aptamer to protein ratio, the band of the free aptamer disappears and the band with lower mobility corresponding to Vap7-NH_2-$VEGF_{165}$ complex appears. Comparing this result with what seen for unmodified Vap7 (Figure 2.31), the 5'-amino modification does not impair the Vap7-VEGF binding. In Figure 2.32b is reported the result obtained using the Vap7-FAM: the shift of the free aptamer starts at the 1:50 aptamer to protein ratio. 5'-FAM modification results to be an important chemical modification that compromises the aptamer binding to VEGF. The planar and hydrophobic surfaces individuated by the fluorophore (5'-FAM modification) do not impair the binding, but indeed they determine a loss of binding affinity.

2.4.2.1.2 VEa5-$VEGF_{165}$ Interaction in Solution

VEa5 is the 66-mer DNA aptamer against $VEGF_{165}$ that bind $VEGF_{165}$ to the heparin binding domain (HBD) with a K_d value of 130 nM [140]. VEa5 was planned to be used in tandem with Vap7 aptamer. VEa5 in fact binds $VEGF_{165}$ to a different epitope from the other aptamer and this different recognition pattern allows the construction of a sandwich assay system.

As for Vap7, the $VEGF_{165}$-binding complex formation was not deeply studied in literature and no information is known about VEa5 adaptability as capture or detection layer in solid phase. Therefore, we investigated, using unmodified VEa5 as control, the effect of both the 5'-amino and the 5'-FAM modifications on protein binding. VEa5, VEa5-NH_2 or VEa5-FAM were separately denatured, folded and incubated with increasing concentrations (up to 10 µM) of $VEGF_{165}$ in TBSE and the complex formation was analyzed by EMSA. The presence of potassium is not required to obtain the correct folding of VEa5 aptamers [140].

In Figure 2.33 are reported the results obtained by EMSA of unmodified VEa5 (a), of VEa5-NH_2 (b) and of VEa5-FAM (c) with $VEGF_{165}$.

Unmodified VEa5, used as control, binds to $VEGF_{165}$ from aptamer/protein rato 1:5 (Figure 2.33a): the band corresponding to VEa5 shifts up and a new band appears in the upper part of the gel, as expected. Even when VEa5-NH_2 (Figure 2.33b) and VEa5-FAM (Figure 2.33c) were analyzed the gel shift was obtained, but the band of free aptamer, in both cases,

shifts up starting from the aptamer/protein ratio 1:50. This means that both the 5'-modifications introduced on VEa5 compromise the binding to $VEGF_{165}$, reducing 10 times the binding affinity. However, the presence of the chemical modifications allows protein recognition, since the binary complex formation is not abolished.

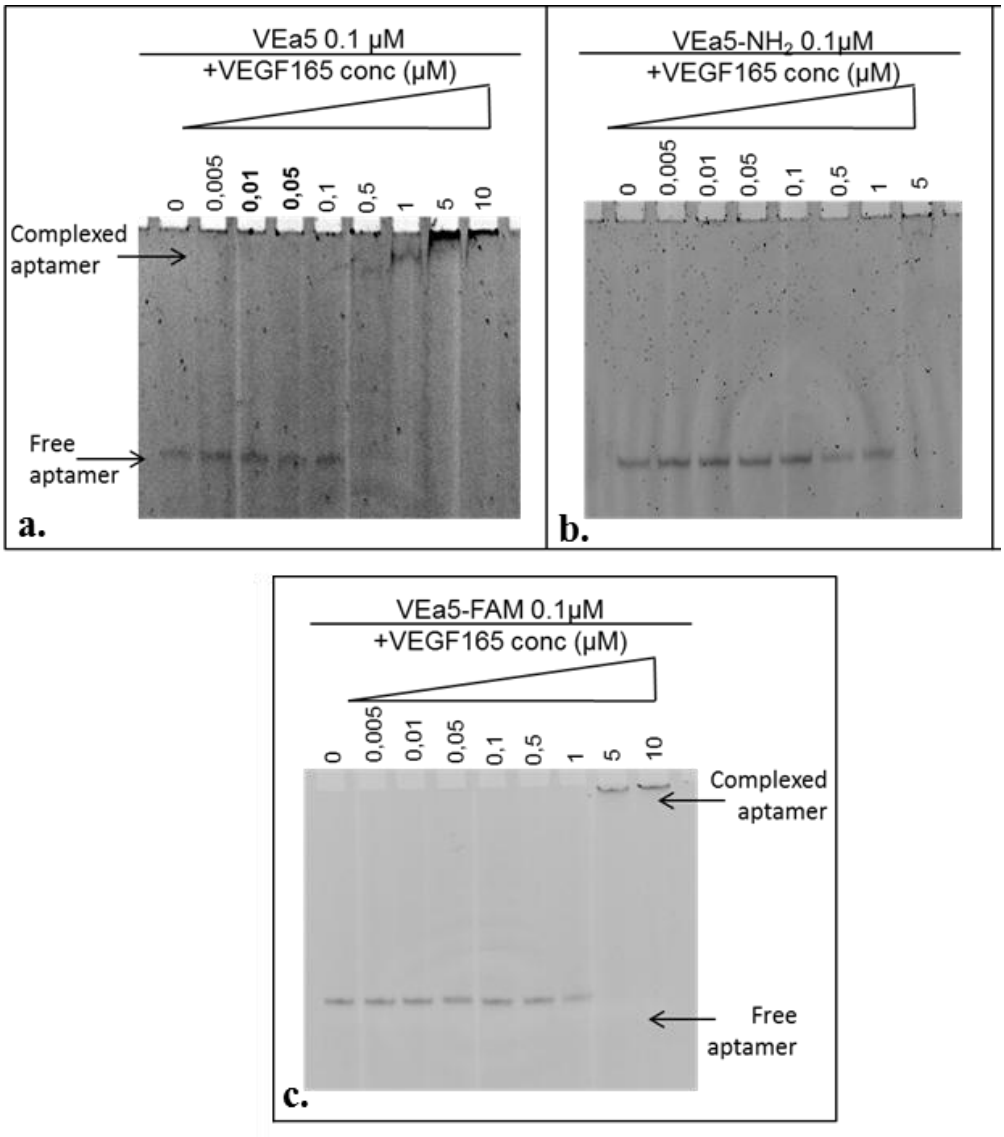


Figure 2.33 Electrophoretic Mobility Shift Assay (EMSA) of unmodified VEa5 with $VEGF_{165}$ (a), of VEa5-NH_2 with $VEGF_{165}$ (b) and of VEa5-FAM with $VEGF_{165}$ (c). VEa5, VEa5-NH_2 and VEa5-FAM were separately incubated with the protein under the conditions described. We have indicated in the figure the respective aptamer and VEGF concentrations. A control reaction without VEGF was performed in all experiments. Binding reactions were applied on a 12% non-denaturing PAA gel containing 1X glycine buffer. The mobility of free and complexed aptamer, stained by SybrGreen II®(a, b), was detected using the Geliance 600 Imaging System.

2.4.2.1.3 Supershift Assay: Optimization of Vap7-VEGF165-VEa5 Interaction

After having completed the studies of binding for each set of modified aptamers, we verified the formation of the ternary complex in solution. The Supershift Assay using the aptamers with the modifications needed for solid phase will indicate the feasibility of the sandwich system in the Sandwich Aptamer Microarray. The Supershift Assay will reveal the best combination of aptamer in order to start the experiments in solid phase. The Supershift Assay was therefore performed with two combinations of aptamers: Vap7-NH_2 as capture aptamer and VEa5-FAM as detection aptamer; VEa5-NH_2 as capture layer and Vap7-FAM as detection layer. The primary aptamer and the secondary aptamer were folded in the suitable buffer, mixed and then incubated with $VEGF_{165}$ in TBSE buffer, at 25°C for 30 min. After incubation, free aptamer and $VEGF_{165}$-aptamer complexes were resolved by 12% non-denaturing polyacrylamide gel containing Glycine buffer 50 mM pH 9.5 and KCl 10 mM. Free aptamers and the binary complex of each aptamer with $VEGF_{165}$ were loaded in the gel as controls. The result of the experiment is reported in Figure 2.34.

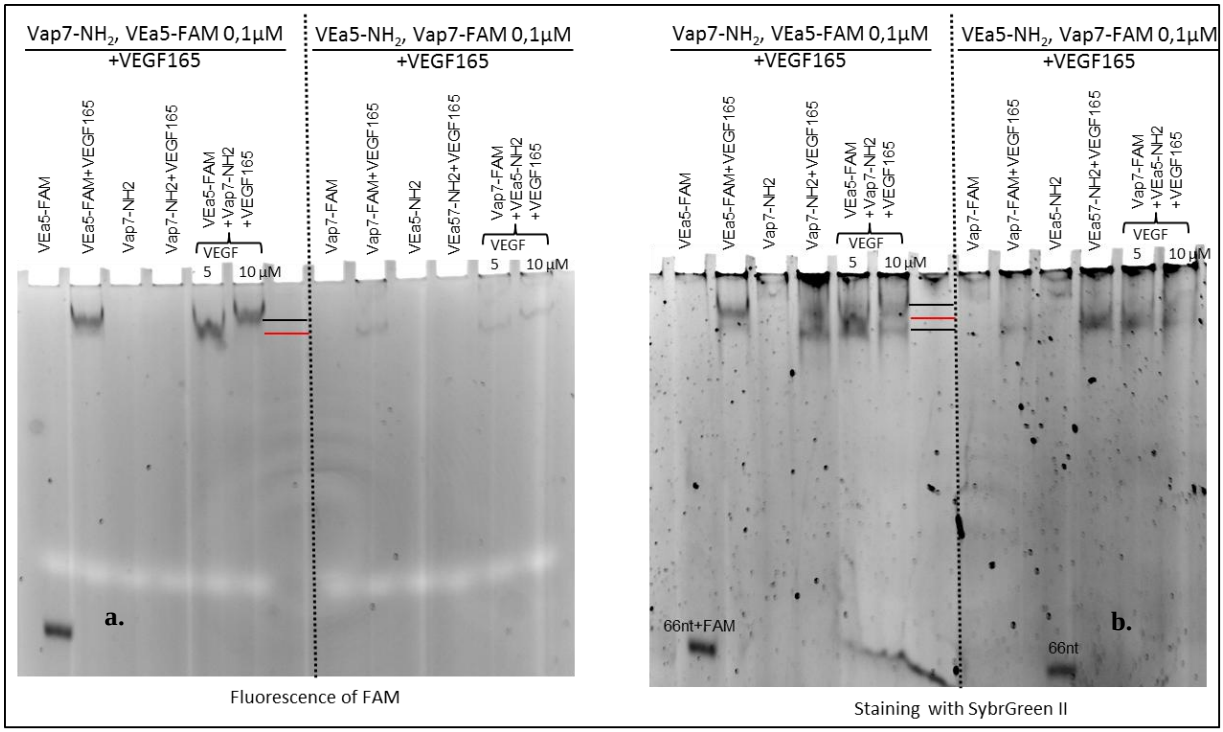


Figure 2.34 Electrophoretic Mobility Supershift Assay (EMSA) of Vap7-NH_2 + $VEGF_{165}$+ VEa5-FAM and of VEa5-NH_2 + $VEGF_{165}$+ Vap7-FAM. Fluorescence of the fluorophore FAM (a) and staining with Sybr Green II (b). On the left part of the gel the Vap7-NH_2-$VEGF_{165}$-VEa5-FAM interaction was analyzed, while on the right part of the gel the VEa5-NH_2-$VEGF_{165}$-Vap7-FAM interaction was analyzed. Each aptamer was incubated separately or simultaneously with the protein under the conditions described. We have indicated in the figure the respective aptamer and $VEGF_{165}$ concentrations. A reaction control without protein was performed for each aptamer. Binding reactions were applied on a 12% non-denaturing PAA gel containing Gycine buffer and KCl 10 mM. The mobility of free and complexed aptamers, stained by SybrGreen II®(b), was detected using the

Geliance 600 Imaging System. On the left part of the gel (both a and b), with the black lines are indicated the positions of the binary complexes; with the red line is indicated the position of the ternary complex.

The detection of the FAM fluorescence signal (Figure 2.34a) allows to identify the shift of the fluoresceinated aptamer (VEa5-FAM or Vap7-FA) when it binds to the protein. Also the fluorescence signal of the ternary complex can be distinguished, since the detection aptamer is involved in the sandwich system.

On the left part of the gel we analyzed Vap7-NH_2-$VEGF_{165}$-VEa5-FAM interaction: at 5μM $VEGF_{165}$ concentration, a new fluorescent band with higher electrophoretic mobility than the binary complex evidently appears. This new band with different hydrodynamic volume can correspond to the ternary complex. This is confirmed by the staining with Sybr Green II (Figure 2.34b): the new band shows a different mobility from both the Vap7-NH_2-$VEGF_{165}$ and the Vea5-FAM-$VEGF_{165}$ binary complexes. At 10 μM $VEGF_{165}$ concentration, the ternary complex results to be loose: the two bands corresponding to the binary complexes reappear, as seen in the gel stained with SybrGreen (Figure 2.34b).

On the right part of the gel we analyzed VEa5-NH_2-$VEGF_{165}$-Vap7-FAM interaction: a sandwich complex is not detectable, since no evident differences from the binary complexes could be observed when VEa5-NH_2 and Vap-FAM were incubated simultaneously with the protein.

Hence, the Supershift Assay indicates that in solution only the Vap7-NH_2-$VEGF_{165}$-VEa5-FAM ternary complex is formed, and therefore this system can be applied to the solid phase: Vap7-NH_2 will be used as capture layer while the VEa5 labeled with the fluorophore as detection layer.

2.4.2.2 Analysis in Solid Phase

With the aim of developing a Sandwich Aptamer Microarray for $VEGF_{165}$ detection, the sandwich system with the modified anti-VEGF aptamers analyzed in solution was applied in solid phase. The pair of aptamer suitable for the SAM resulted to be Vap7 as capture layer and VEa5 as detection layer (Figure 2.34).

2.4.2.2.1 Optimization of the Aptamer-$VEGF_{165}$ Sandwich Protocol

The first goal of the analysis in solid phase was that to improve the $VEGF_{165}$ capture by the capture aptamer. As a result of what was learned from the development of the SAM for thrombin detection, we anchored on the glass slide not only the 5'-amino modified Vap7 aptamer (Vap7-NH_2), but also the 5'-amino modified plus a polyT (12T) tail as spacer

[Vap7(12T)NH_2], for making the G-quadruplex motif of the aptamer, responsible of $VEGF_{165}$ recognition, more accessible to the target protein. Alexa555-labeled $VEGF_{165}$ was used to detect the captured protein in correspondence of the two different aptamers. In Figure 2.35 is reported the result of this experiment.

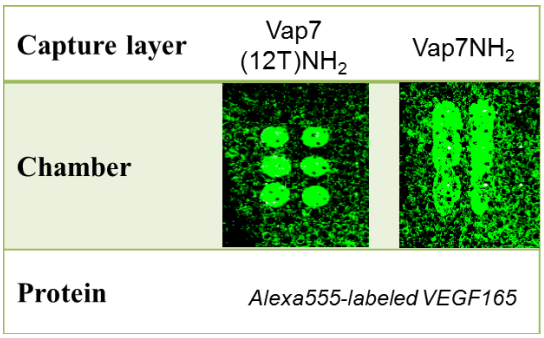


Figure 2.35 Images of the slide after the incubation of Alexa555-labeled $VEGF_{165}$. On the left Vap7(12T)NH_2 was anchored on the glass slide as capture layer; on the right Vap7-NH_2 was anchored on the glass slide as capture layer. Alexa555 emitted green light and for this reason green fluorescence represents labeled protein.

Comparing the two chambers in which Vap7(12T)NH_2 (on the left) and Vap7-NH_2 (on the right) were anchored, we can observe that the fluorescence signal of the Alexa555-labeled $VEGF_{165}$ results evidently to be better defined and sharp in correspondence of Vap7(12T)NH_2. Moreover, in the chamber on the left, the background signal results to be less intense, probably because the fluorescent protein is well captured and concentrated on the capture aptamer. For this reason we considered as optimal capture layer the aptamer Vap7(12T)NH_2 and it was printed on the microarray slides for the following analysis in solid phase.

The SAM was then performed employing Vap7(12T)NH_2 as capture layer and the fluorescently labeled VEa5 as detection layer. As resulted in the case of thrombin detection by SAM, the fluorescent signal of FAM was not detected, probably due to a limit of detection of our laser scanner; therefore, we decided to perform our SAM for $VEGF_{165}$ detection using the 5'-Cy5 modified VEa5 aptamer (VEa5-Cy5). In this case sandwich formation should be detectable by red fluorescence of VEa5-Cy5.

Alexa555-labeled $VEGF_{165}$ was also used as control to further verify the sandwich formation by the yellow fluorescence resulting from the co-localization of green (Alexa conjugated to the protein) and red (Cy conjugated to the oligonucleotide) fluorophores. Both the one-step and the two-steps protocols previously described were tested for the sandwich formation on surface. The slides carefully rinsed with the PBS Buffer and MilliQ H_2O were

finally spin-dried and scanned. The samples (six spots for each sample) were prepared as shown schematically in Table 2.7; each sub-array was physically isolated from the others.

Sub-array	1		2		3		4	
Capture layer	Vap7 (12T)NH_2	TBA1 (12T)NH_2	Vap7 (12T)NH_2	TBA1 (12T)NH_2	Vap7 (12T)NH_2	TBA1 (12T)NH_2	Vap7 (12T)NH_2	TBA1 (12T)NH_2
Protein	$VEGF_{165}$ (0,1µM) pre-complexed with		Alexa555-$VEGF_{165}$ (0,1µM) pre-complexed with		$VEGF_{165}$ (0,1µM)		Alexa555-$VEGF_{165}$ (0,1µM)	
Detection layer	VEa5-Cy5(0,5µM)		VEa5-Cy5(0,5µM)		VEa5-Cy5(0,5µM)		VEa5-Cy5(0,5µM)	
Procedure	One-step				Two-steps			

Table 2.7 Schematic representation of the samples incubated in the SAM protocol.

The results of the SAM experiments are shown in Figure 2.36.

In chambers 1 the red spots (VEa5-Cy5) in correspondence of the selection aptamer Vap7(12T)NH_2 anchored on the glass slide reveal that the sandwich was formed with specificity for the anti-$VEGF_{165}$ aptamer: the detection of the Cy5 labeled aptamer is reproducible across the six spots of each sub-array. The recognition is also specific: no signal is detected for TBA1(12T)NH_2.

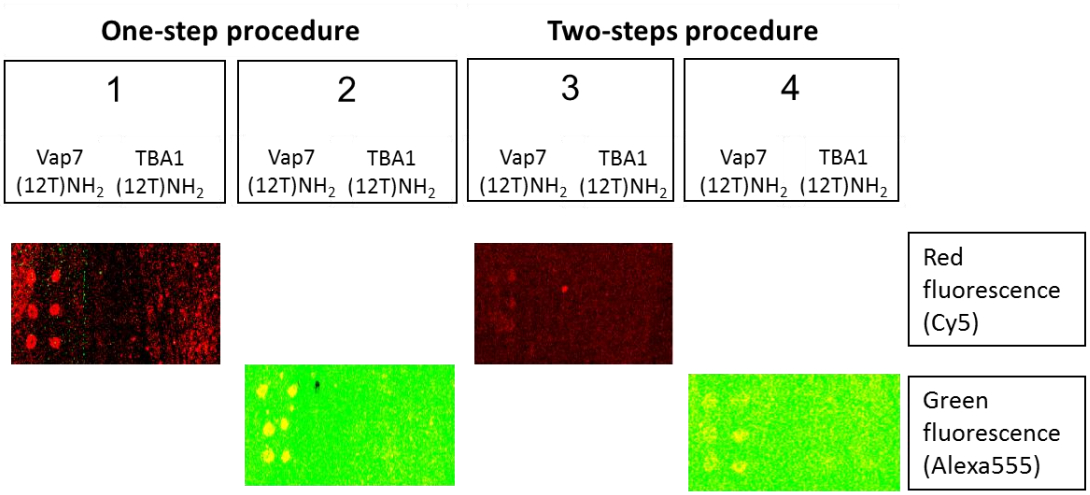


Figure 2.36 Images of the slide after the Sandwich Aptamer Microarray (SAM). In each subarray Vap7(12T)NH_2 was printed on the left, while TBA1(12T)NH_2 on the right. The fluorophore Cy5 emitted red light, while Alexa555 emitted green light. For this reason red fluorescence represents VEa5-Cy5 and green fluorescence represents labeled $VEGF_{165}$. In chambers 2 and 4, in which labeled protein was incubated, the fluorescence was detected with the contrast to reveal green light: yellow spots indicate the simultaneously localization of the two fluorophores (red and green).

SAM is formed with $VEGF_{165}$ and also with Alexa-labeled $VEGF_{165}$: in the sub-arrays 2, Alexa-labeled protein allows a further control for $VEGF_{165}$ presence in the sandwich by the yellow fluorescence resulting from the simultaneous presence of green fluorescence (Alexa555-$VEGF_{165}$) and red fluorescence (VEa5-Cy5). The one-step protocol is more efficient than the two step procedure: the detection of the sandwich is sharp and specific in chambers 1 and 2 compared with chambers 3 and 4. In chamber 3 the signal obtained is weak, while chamber 4 (Alexa-labeled protein) exhibits lower selectivity. The one-step protocols ensures enough time to the modified VEa5 aptamer for the formation of the correct binary complex able to expose the protein epitopes recognized by the immobilized capture aptamer. According with the analysis in solution, the presence of a 5'-modification has a negative effect on VEa5, possibly influencing the pattern mode of recognition VEa5-$VEGF_{165}$, reflected in the loss of specificity of binding observed in solid phase.

The microarray system protocol of chamber 1 results a successful application of the aptamer sandwich devised.

2.4.2.3 Specificity of protein recognition

SAM resulted a successful application both for thrombin and for $VEGF_{165}$ detection, separately. In view of future applications in multiplex microarray systems, the two proteins, Alexa555-labeled $VEGF_{165}$ and Alexa647-labeled thrombin, were mixed and incubated simultaneously in a chamber in which Vap7(12T)NH_2 on the left and TBA1(12T)NH_2 on the right were printed on the glass slide as capture layers. The result of the experiment is reported in Figure 2.37.

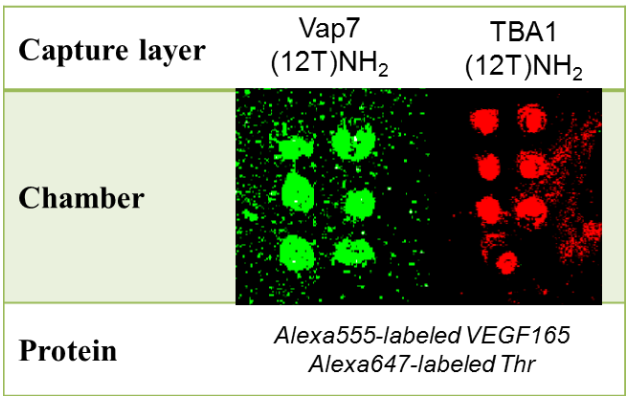


Figure 2.37 Images of the slide after the simultaneous incubation of Alexa647-labeled thrombin and Alexa555-labeled $VEGF_{165}$. On the left Vap7(12T)NH_2 was anchored on the glass slide as capture layer for $VEGF_{165}$; on the right TBA1(12T)NH_2 was anchored on the glass slide as capture layer for thrombin. Alexa555 emitted green light and Alexa647 emitted red light, for this reason green fluorescence represents labeled $VEGF_{165}$ and red fluorescence represents labeled thrombin.

Alexa555 emitted green light and Alexa647 emitted red light, so green fluorescence represents labeled $VEGF_{165}$ and red fluorescence represents labeled thrombin immobilized on the glass slide by the specific aptamer. The green spots of Figure 2.37 in correspondence of Vap7(12T)NH_2 denote the presence of $VEGF_{165}$, while the red spots in correspondence of TBA1(12T)NH_2 indicate the capture of thrombin. Hence, each protein, also in a complex mix with the other one, is specifically recognized by its proper capture aptamer.

This experiment represents the starting point for the development of SAM as potential tool for the simultaneous diagnosis of different diseases-related target protein.

2.5 CONCLUSIONS

The present study reports the development of a Sandwich Aptamer Microarray (SAM) for protein detection using DNA aptamers modified at the 5'-end for the application to the surface of glass slides and for fluorescence detection.

The first goal achieved in this work is the development of the SAM system for human thrombin detection. An extensive analysis, optimization and validation of binary and ternary complexes formation between the modified aptamers and thrombin in solution was followed by the verification of the aptamer sandwich formation in solid phase. Our results add information on the effects of modifying aptamers for microarray applications: the 5'-modifications to TBA1 have a minimal but observable effect on thrombin recognition in solution. In solid phase the aptamer orientation through the polyT spacer allows for the correct presentation to the protein, in agreement with the results obtained by other authors analyzing in solid phase the binding of 5'-end modified aptamers to Cy-labeled thrombin [121]. Differently from the systems adopted by these authors, we resorted to a classical sandwich for detecting thrombin. The labeling of TBA2 with the fluorophore Cy5 influences the complex formation with thrombin. Our results however show that the binding to the protein target is preserved in solution when TBA2 is conjugated to Cy5 that was therefore adopted for fluorescent detection. Complex formation was verified analyzing the glass slides by a microarray laser scanner, and correct sandwich formation was confirmed by the fluorescence signals of the sandwich (Figure 2.25). The microarray for thrombin detection indeed results specific and efficient, therefore representing a potential tool for diagnostic application.

The SAM system for thrombin detection was further characterized and optimized to be developed as a simple and sensitive thrombin diagnostic tool with comparable sensitivity and specificity exhibited by known systems [121, 122, 154, 155]. Analytical approaches based on biosensors for protein recognition and quantification have extensively been analyzed in the past few years [156]. The identification of disease-related biomolecules in body fluids is useful for the diagnosis of diseases as well as for follow-up protocols in clinical settings. Analytes can be identified with different transducing and recognition strategies, including antibodies, enzymes or aptamers: each different approach needs to be optimized in order to achieve increasing selectivity and specificity, and reduced assay time and costs. The obtained data demonstrate that our SAM system based on aptamer recognition of an analyte is efficient and specific, and different detection aptamers and methods can be utilized. The use of the

indirect method (TBA2- biotin plus Cy3-streptavidin) allows lower LOD and LOQ when compared to TBA2-Cy5 (direct method) even in the presence of biological fluids. Moreover, the signal amplification made possible by the use of TBA2-biotin plus Cy3-streptavidin as detection strategy allows for a stronger signal at low protein concentrations. The SAM exhibits a limit of detection comparable or superior to other systems described in the literature employing the same aptamers and different technologies [157, 158] but has the advantage of a simple set up. Besides, the aptasensor could be further developed with signal amplification strategies by the use of fluorescent nanoparticles conjugated with streptavidin to allow for nanotech systems, to reach lower detection limits comparable to those achieved by advanced detection technologies [123, 159].

In view of future applications of SAM in multiplex microarray systems, we enlarged the SAM system analysis for the $VEGF_{165}$ detection.

Following the same approach adopted for thrombin, a thorough analysis and optimization of binary and ternary complexes formation between the modified anti-$VEGF_{165}$ aptamers and $VEGF_{165}$ in solution was performed prior to verify the sandwich aptamer formation in solid phase. Literature data concerning Vap7 and VEa5 aptamers and their interaction with $VEGF_{165}$ are weak and no detection system different from ELISA (VEGF Human ELISA Kit, Life Technologies) are reported. Our pioneering work about $VEGF_{165}$ detection through the aptamer-based microarray provides new knowledge about Vap7 and VEa5 binding to $VEGF_{165}$. The optimal binding conditions were identified and the effects of 5'-chemical modifications to Vap7 and VEa5 aptamers for microarray applications were investigated. The presence of a labeling dye conjugated to both the Vap7 and the VEa5 aptamers allows $VEGF_{165}$ recognition, since the binary complex formation is not abolished, but it induces a decrease in binding affinity. The 5'-amino modification compromises only the binding of VEa5 to the protein. The solution experiments point to the requirement of careful handling of anti-$VEGF_{165}$ aptamers when adapting the solution studies to the solid phase. The pair of modified aptamers that ensures the sandwich formation was identified in solution and the system was then applied in solid phase. To ensure $VEGF_{165}$ recognition, Vap7(12T)NH_2 was anchored on the glass slides as capture layer for making the G-quadruplex motif more accessible to the target protein. Consistently with SAM for thrombin detection, VEa5 conjugated to Cy5 was adopted for fluorescent detection of $VEGF_{165}$. The results obtained in the microarray are positive and consistent with the results obtained by the analysis in solution,

and constitute the proof of principle of our approach of validating in solution the effects of chemical modification of aptamers prior or along the development of the solid phase.

Hence, the sandwich aptamer formation was verified also for $VEGF_{165}$ and its aptamers analyzing the glass slides by a microarray laser scanner, and correct sandwich formation was confirmed by the fluorescence signals of the sandwich (Figure 2.36).

Further characterization for the limit of detection and optimization for the best incubation time, temperature and serum presence by biological samples of the solid phase microarray system for $VEGF_{165}$ detection are in progress.

Moreover, preliminary studies suggest that SAM system could be exploited for the development of a multiplex microarray for protein analysis in presence of a single detection strategy, *i.e.* for the simultaneous detection of different target proteins.

The data have been published in:

Sosic, A., Meneghello, A., Cretaio, E. and Gatto, B. *Human thrombin detection through a sandwich aptamer microarray: interaction analysis in solution and in solid phase.* Sensors (Basel), 2011. **11**(10): p. 9426-41.

Meneghello, A., Sosic, A., Antognoli, A., Cretaio, E. and Gatto, B. *Development and optimization of a thrombin Sandwich Aptamer Microarray*. Microarrays, 2012. **1**: p. 95-106.

ABBREVIATION LIST

Ab antibody

Alexa Alexa555® fluorophore

ALISA aptamer-linked immobilized sorbent assay

AMD age-related macular degeneration

AuNPs Au nanoparticles

AuNRs Au nanorods

CD circular dichroism

CNTs carbon nanotubes

CV cyclic voltammetry

Cy5 Cyanine dye

cTAR complementary Trans-Activation responsive region

DAB dabcyl, 4-(4'-dimethylaminophenylazo)benzoic acid

DIBA disulfide benzamide

DMSO Dimethyl sulfoxide

DPV differential pulse voltammetry

dsDNA double strand DNA

EA ellagic acid

ELISA enzyme-linked immunosorbent assay

ELs ellagic derivatives

EMSA Electrophoretic Mobility Shift Assay

EtBr GelRed^{TM} Nucleic Acid Gel Stain

ETN Tris HCl + NaCl + EDTA buffer

FAM carboxyfluorescein

FBD Fibrinogen Binding Domain

FBS Fetal Bovine Serum

Fc ferrocene

FRET Förster Resonance Energy Transfer

FQA Fluorescence Quenching Assay

GA gallic acid

GLB gel loading buffer

G-4 G-quadruplex

GO grapheme oxide

HAART highly active antiretroviral therapy

HBD Heparin Binding Domain

HTS High Throughput Screening

Kd constant of dissociation

Ki constant of inhibition

ICP-MS inductively coupled plasma mass spectrometry

IC_{50} half maximal inhibitory concentration

LOD limit of detection

LOQ limit of quantification

LTR long terminal repeat region

Mab monoclonal antibody

MB methylene blue

NAS *N*,*N*-acryloyloxysuccinimide

NC nucleocapsid protein

NOBA 3-nitroso benzamide

OTA Ochratoxin A binding aptamer

PAA polyacrylamide

PBS primer binding site

PIC Peptide-Intercalator Conjugate

POCT point-of-care testing

PPT poly-purine tract

QD quantum dots

RBD receptor binding domain

RDT rapid diagnostic testing

RT reverse transcriptase

SAM Sandwich Aptamer Microarray

SELEX Systematic Evolution of Ligands by EXponential enrichment

SDS Sodium Dodecyl Sulfate

SPR surface plasmon resonance

SWNTs single wall nanotubes

TAR Trans-Activation Responsive element

Tat transactivator of transcription

TBA1 unmodified 15-mer Thrombin Binding Aptamer 1

TBA1-NH_2 5'-amino modified Thrombin Binding Aptamer 1

TBA1(12T)NH_2 5'-amino modified plus a polyT(12) spacer Thrombin Binding Aptamer 1

TBA2 unmodified 29-mer Thrombin Binding Aptamer 2

TBA2-Alexa 5'-AlexaFluor488 modified Thrombin Binding Aptamer 2

TBA2Cy5 5'-Cy5 modified Thrombin Binding Aptamer 2

TBE Tris HCl + Boric acid + EDTA buffer

Thr Thrombin

Tm melting temperature

TN Tris HCl + NaCl buffer

TNMg Tris HCl + NaCl + $Mg(ClO_4)_2$ buffer

UA urolithin A

UB urolithin B

Vap7 unmodified Vap7 anti-VEGF aptamer

Vap7* [γ^{32}P]ATP-labeled Vap7

Vap7-FAM 5'-FAM modified Vap7 anti-VEGF aptamer

Vap7-NH_2 5'-amino modified Vap7 anti-VEGF aptamer

Vap7(12T)NH_2 5'-amino modified plus a polyT(12) spacer Vap7 anti-VEGF aptamer

VEa5 unmodified VEa5 anti-VEGF aptamer

VEa5-FAM 5'-FAM modified VEa5 anti-VEGF aptamer

VEa5-Cy5 5'-Cy5 modified VEa5 anti-VEGF aptamer

VEa5-NH_2 5'-amino modified VEa5 anti-VEGF aptamer

VEGF vascular endothelial growth factor

ZF Zinc finger

6-DFQs 6-desfluoroquinolones

(-)SSDNA minus strong stop DNA

REFERENCES

1. Pace, P. and M. Rowley, *Integrase inhibitors for the treatment of HIV infection.* Curr Opin Drug Discov Devel, 2008. **11**(4): p. 471-9.
2. Van Heuverswyn, F., et al., *Human immunodeficiency viruses: SIV infection in wild gorillas.* Nature, 2006. **444**(7116): p. 164.
3. Taylor, B.S. and S.M. Hammer, *The challenge of HIV-1 subtype diversity.* N Engl J Med, 2008. **359**(18): p. 1965-6.
4. Bohnlein, S., J. Hauber, and B.R. Cullen, *Identification of a U5-specific sequence required for efficient polyadenylation within the human immunodeficiency virus long terminal repeat.* J Virol, 1989. **63**(1): p. 421-4.
5. Feng, S. and E.C. Holland, *HIV-1 tat trans-activation requires the loop sequence within tar.* Nature, 1988. **334**(6178): p. 165-7.
6. Wu, Y., *HIV-1 gene expression: lessons from provirus and non-integrated DNA.* Retrovirology, 2004. **1**: p. 13.
7. Evering, T.H. and M. Markowitz, *Raltegravir: an integrase inhibitor for HIV-1.* Expert Opin Investig Drugs, 2008. **17**(3): p. 413-22.
8. Weidner, N., *Intratumor microvessel density as a prognostic factor in cancer.* Am J Pathol, 1995. **147**(1): p. 9-19.
9. Cheryn, P., *HIV treatments and highly active antiretroviral therapy.* Australian Prescriber, 2003. **26-3**.
10. Hsiao-Wei Liu, B.S.C., *Single-Molecule Studies on the Role of HIV-1 NucleocapsidProtein/Nucleic Acid Interaction in the Viral Replication Cycle*. 2007, University of Texas Libraries, Austin.
11. Berg, J.M., *Potential metal-binding domains in nucleic acid binding proteins.* Science, 1986. **232**(4749): p. 485-7.
12. Covey, S.N., *Amino acid sequence homology in gag region of reverse transcribing elements and the coat protein gene of cauliflower mosaic virus.* Nucleic Acids Res, 1986. **14**(2): p. 623-33.
13. Green, L.M. and J.M. Berg, *Retroviral nucleocapsid protein-metal ion interactions: folding and sequence variants.* Proc Natl Acad Sci U S A, 1990. **87**(16): p. 6403-7.
14. Henderson, L.E., et al., *Primary structure of the low molecular weight nucleic acid-binding proteins of murine leukemia viruses.* J Biol Chem, 1981. **256**(16): p. 8400-6.
15. Darlix, J.L., et al., *First glimpses at structure-function relationships of the nucleocapsid protein of retroviruses.* J Mol Biol, 1995. **254**(4): p. 523-37.
16. Turner, B.G. and M.F. Summers, *Structural biology of HIV.* J Mol Biol, 1999. **285**(1): p. 1-32.
17. Tanchou, V., et al., *Formation of stable and functional HIV-1 nucleoprotein complexes in vitro.* J Mol Biol, 1995. **252**(5): p. 563-71.
18. Stote, R.H., et al., *Structure of the His44 --> Ala single point mutant of the distal finger motif of HIV-1 nucleocapsid protein: a combined NMR, molecular dynamics simulation, and fluorescence study.* Biochemistry, 2004. **43**(24): p. 7687-97.
19. Stoylov, S.P., et al., *Ordered aggregation of ribonucleic acids by the human immunodeficiency virus type 1 nucleocapsid protein.* Biopolymers, 1997. **41**(3): p. 301-12.
20. Le Cam, E., et al., *Properties and growth mechanism of the ordered aggregation of a model RNA by the HIV-1 nucleocapsid protein: an electron microscopy investigation.* Biopolymers, 1998. **45**(3): p. 217-29.
21. Fisher, R.J., et al., *Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides.* J Virol, 1998. **72**(3): p. 1902-9.
22. Levin, J.G., et al., *Nucleic acid chaperone activity of HIV-1 nucleocapsid protein: critical role in reverse transcription and molecular mechanism.* Prog Nucleic Acid Res Mol Biol, 2005. **80**: p. 217-86.

23. Bernacchi, S., et al., *HIV-1 nucleocapsid protein activates transient melting of least stable parts of the secondary structure of TAR and its complementary sequence.* J Mol Biol, 2002. **317**(3): p. 385-99.
24. Beltz, H., et al., *Structural determinants of HIV-1 nucleocapsid protein for cTAR DNA binding and destabilization, and correlation with inhibition of self-primed DNA synthesis.* J Mol Biol, 2005. **348**(5): p. 1113-26.
25. Rein, A., *Nucleic acid chaperone activity of retroviral Gag proteins.* RNA Biol, 2010. **7**(6): p. 700-5.
26. Godet, J. and Y. Mely, *Biophysical studies of the nucleic acid chaperone properties of the HIV-1 nucleocapsid protein.* RNA Biol, 2010. **7**(6): p. 687-99.
27. Kanevsky, I., et al., *Structural determinants of TAR RNA-DNA annealing in the absence and presence of HIV-1 nucleocapsid protein.* Nucleic Acids Res, 2011. **39**(18): p. 8148-62.
28. Druillennec, S., et al., *A mimic of HIV-1 nucleocapsid protein impairs reverse transcription and displays antiviral activity.* Proc Natl Acad Sci U S A, 1999. **96**(9): p. 4886-91.
29. Kim, M.Y. and S. Jeong, *Inhibition of the functions of the nucleocapsid protein of human immunodeficiency virus-1 by an RNA aptamer.* Biochem Biophys Res Commun, 2004. **320**(4): p. 1181-6.
30. de Rocquigny, H., et al., *Targeting the viral nucleocapsid protein in anti-HIV-1 therapy.* Mini Rev Med Chem, 2008. **8**(1): p. 24-35.
31. Stephen, A.G., et al., *Identification of HIV-1 nucleocapsid protein: nucleic acid antagonists with cellular anti-HIV activity.* Biochem Biophys Res Commun, 2002. **296**(5): p. 1228-37.
32. Davis, W.R., et al., *Actinomycin D inhibition of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase and nucleocapsid protein.* Biochemistry, 1998. **37**(40): p. 14213-21.
33. Godet, J., et al., *During the early phase of HIV-1 DNA synthesis, nucleocapsid protein directs hybridization of the TAR complementary sequences via the ends of their double-stranded stem.* J Mol Biol, 2006. **356**(5): p. 1180-92.
34. Nikolov, D.B. and S.K. Burley, *RNA polymerase II transcription initiation: a structural view.* Proc Natl Acad Sci U S A, 1997. **94**(1): p. 15-22.
35. Zhu, Y., et al., *Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation in vitro.* Genes Dev, 1997. **11**(20): p. 2622-32.
36. Wei, P., et al., *A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA.* Cell, 1998. **92**(4): p. 451-62.
37. Bieniasz, P.D., et al., *Recruitment of cyclin T1/P-TEFb to an HIV type 1 long terminal repeat promoter proximal RNA target is both necessary and sufficient for full activation of transcription.* Proc Natl Acad Sci U S A, 1999. **96**(14): p. 7791-6.
38. Neuveut, C. and K.T. Jeang, *Recombinant human immunodeficiency virus type 1 genomes with tat unconstrained by overlapping reading frames reveal residues in Tat important for replication in tissue culture.* J Virol, 1996. **70**(8): p. 5572-81.
39. Wang, Z. and T.M. Rana, *RNA conformation in the Tat-TAR complex determined by site-specific photo-cross-linking.* Biochemistry, 1996. **35**(20): p. 6491-9.
40. Wang, Z. and T.M. Rana, *RNA-protein interactions in the Tat-trans-activation response element complex determined by site-specific photo-cross-linking.* Biochemistry, 1998. **37**(12): p. 4235-43.
41. Churcher, M.J., et al., *High affinity binding of TAR RNA by the human immunodeficiency virus type-1 tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region.* J Mol Biol, 1993. **230**(1): p. 90-110.
42. Tao, J. and A.D. Frankel, *Electrostatic interactions modulate the RNA-binding and transactivation specificities of the human immunodeficiency virus and simian immunodeficiency virus Tat proteins.* Proc Natl Acad Sci U S A, 1993. **90**(4): p. 1571-5.
43. Aboul-ela, F., J. Karn, and G. Varani, *Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge.* Nucleic Acids Res, 1996. **24**(20): p. 3974-81.

44. Kashanchi, F., M.R. Sadaie, and J.N. Brady, *Inhibition of HIV-1 transcription and virus replication using soluble Tat peptide analogs.* Virology, 1997. **227**(2): p. 431-8.
45. Martinez, M.A., et al., *Suppression of chemokine receptor expression by RNA interference allows for inhibition of HIV-1 replication.* AIDS, 2002. **16**(18): p. 2385-90.
46. Thomas, J.R. and P.J. Hergenrother, *Targeting RNA with small molecules.* Chem Rev, 2008. **108**(4): p. 1171-224.
47. Tabarrini, O., et al., *Studies of anti-HIV transcription inhibitor quinolones: identification of potent N1-vinyl derivatives.* ChemMedChem, 2010. **5**(11): p. 1880-92.
48. Sissi, C., et al., *Mg(2+)-mediated binding of 6-substituted quinolones to DNA: relevance to biological activity.* Bioorg Med Chem, 1998. **6**(9): p. 1555-61.
49. Manfroni, G., et al., *Synthesis and biological evaluation of 2-phenylquinolones targeted at Tat/TAR recognition.* Bioorg Med Chem Lett, 2009. **19**(3): p. 714-7.
50. Giaretta, G., *Acidi nucleici come target terapeutici nello sviluppo di composti antivirali e antiproliferativi*, in *Dipartimento di Scienze Farmaceutiche*. 2008, Università degli Studi di Padova: Padova.
51. Furlanetto, V., et al., *Ellagic acid and polyhydroxylated urolithins are potent catalytic inhibitors of human topoisomerase II: an in vitro study.* J Agric Food Chem, 2012. **60**(36): p. 9162-70.
52. Quideau, S. and K.S. Feldman, *Ellagitannin Chemistry.* Chem Rev, 1996. **96**(1): p. 475-504.
53. Guo, J., et al., *Human immunodeficiency virus type 1 nucleocapsid protein promotes efficient strand transfer and specific viral DNA synthesis by inhibiting TAR-dependent self-priming from minus-strand strong-stop DNA.* J Virol, 1997. **71**(7): p. 5178-88.
54. Heath, M.J., et al., *Differing Roles of the N- and C-terminal Zinc Fingers in Human Immunodeficiency Virus Nucleocapsid Protein-enhanced Nucleic Acid Annealing.* Journal of Biological Chemistry, 2003. **278**(33): p. 30755-30763 %R 10.1074/jbc.M303819200.
55. de Paula, Q.A., J.B. Mangrum, and N.P. Farrell, *Zinc finger proteins as templates for metal ion exchange: Substitution effects on the C-finger of HIV nucleocapsid NCp7 using M(chelate) species (M=Pt, Pd, Au).* J Inorg Biochem, 2009. **103**(10): p. 1347-54.
56. Omichinski, J.G., et al., *Structural characterization of a 39-residue synthetic peptide containing the two zinc binding domains from the HIV-1 p7 nucleocapsid protein by CD and NMR spectroscopy.* FEBS Lett, 1991. **292**(1-2): p. 25-30.
57. Levin, J.G., et al., *Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription.* RNA Biol. **7**(6): p. 754-74.
58. Rupcich, N., et al., *Quenching of fluorophore-labeled DNA oligonucleotides by divalent metal ions: implications for selection, design, and applications of signaling aptamers and signaling deoxyribozymes.* J Am Chem Soc, 2006. **128**(3): p. 780-90.
59. Dyson, H.J. and P.E. Wright, *Intrinsically unstructured proteins and their functions.* Nat Rev Mol Cell Biol, 2005. **6**(3): p. 197-208.
60. Kanevsky, I., et al., *Structural determinants of TAR RNA-DNA annealing in the absence and presence of HIV-1 nucleocapsid protein.* Nucleic Acids Res. **39**(18): p. 8148-62.
61. Shvadchak, V., et al., *Identification by high throughput screening of small compounds inhibiting the nucleic acid destabilization activity of the HIV-1 nucleocapsid protein.* Biochimie, 2009. **91**(7): p. 916-23.
62. Alcaro, S.G., B.; Giaretta, G.; Bertuzzi, M.; Bianchini, E.; Zagotto, G.; Palumbo, M., *RNA as drug target*, in *New Perspectives In Medicinal Cemistry*, R. Signpost, Editor. 2009, Kerala: India. p. 199-211.
63. Darlix, J.L., et al., *Properties, functions, and drug targeting of the multifunctional nucleocapsid protein of the human immunodeficiency virus.* Adv Pharmacol, 2007. **55**: p. 299-346.
64. Darlix, J.L., et al., *Flexible nature and specific functions of the HIV-1 nucleocapsid protein.* J Mol Biol, 2011. **410**(4): p. 565-81.

65. Sinigaglia, L., *Identification of small molecules inhibiting hiv-1 nc chaperone activity through different mechanisms of action* in *Dipartimento di Scienze del Farmaco*. 2012, Università degli Studi di Padova: Padova.
66. Turner, K.B., et al., *Structural probing of the HIV-1 polypurine tract RNA:DNA hybrid using classic nucleic acid ligands.* Nucleic Acids Res, 2008. **36**(8): p. 2799-810.
67. Sosic, A., ; Casatti, M.; Santagada, V.; Palumbo M.; Sissi, C. and Gatto, B., *Selective recognition of nucleic acids by pepdidyl-anthraquinones*. 2011: Poster comunication at NPCF V Trieste.
68. Cecchetti, V., et al., *6-Aminoquinolones as new potential anti-HIV agents.* J Med Chem, 2000. **43**(20): p. 3799-802.
69. Tabarrini, O., et al., *Structure modifications of 6-aminoquinolones with potent anti-HIV activity.* J Med Chem, 2004. **47**(22): p. 5567-78.
70. Tabarrini, O., et al., *Structure-activity relationship study on anti-HIV 6-desfluoroquinolones.* J Med Chem, 2008. **51**(17): p. 5454-8.
71. Massari, S., et al., *Studies on anti-HIV quinolones: new insights on the C-6 position.* Bioorg Med Chem, 2009. **17**(2): p. 667-74.
72. Furlanetto, V., *Poly-hydroxylated Ellagic Acid and hybrid Resveratrol derivatves as catalytic inhibitors of human and bacterial topoisomerases II*, in *Dipartimento di Scienze del Farmaco*. 2012, Università degli Studi di Padova: Padova.
73. Tuerk, C. and L. Gold, *Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase.* Science, 1990. **249**(4968): p. 505-10.
74. Ellington, A.D. and J.W. Szostak, *In vitro selection of RNA molecules that bind specific ligands.* Nature, 1990. **346**(6287): p. 818-22.
75. Han, K., Z. Liang, and N. Zhou, *Design strategies for aptamer-based biosensors.* Sensors (Basel), 2010. **10**(5): p. 4541-57.
76. Pendergrast, P.S., et al., *Nucleic acid aptamers for target validation and therapeutic applications.* J Biomol Tech, 2005. **16**(3): p. 224-34.
77. Song, K.M., S. Lee, and C. Ban, *Aptamers and their biological applications.* Sensors (Basel), 2012. **12**(1): p. 612-31.
78. Mascini, M., *Aptamers and their applications.* Anal Bioanal Chem, 2008. **390**(4): p. 987-8.
79. Birch, J.R. and A.J. Racher, *Antibody production.* Adv Drug Deliv Rev, 2006. **58**(5-6): p. 671-85.
80. Jayasena, S.D., *Aptamers: an emerging class of molecules that rival antibodies in diagnostics.* Clin Chem, 1999. **45**(9): p. 1628-50.
81. Ireson, C.R. and L.R. Kelland, *Discovery and development of anticancer aptamers.* Mol Cancer Ther, 2006. **5**(12): p. 2957-62.
82. Bunka, D.H. and P.G. Stockley, *Aptamers come of age - at last.* Nat Rev Microbiol, 2006. **4**(8): p. 588-96.
83. Mohanty, S.P.a.K., E., *Biosensors: A tutorial review.* IEEE potentials, 2006.
84. Strehlitz, B.N., N.; Stoltenburg, R., *Protein Detection with Aptamer Biosensors.* Sensors, 2008. **8**: p. 4296-4307.
85. Rahman, M.A., et al., *Gold nanoparticles doped conducting polymer nanorod electrodes: ferrocene catalyzed aptamer-based thrombin immunosensor.* Anal Chem, 2009. **81**(16): p. 6604-11.
86. Wang, X., et al., *Detection of thrombin using electrogenerated chemiluminescence based on Ru(bpy)3(2+)-doped silica nanoparticle aptasensor via target protein-induced strand displacement.* Anal Chim Acta, 2007. **598**(2): p. 242-8.
87. Chandra, P., et al., *Detection of daunomycin using phosphatidylserine and aptamer co-immobilized on Au nanoparticles deposited conducting polymer.* Biosens Bioelectron, 2011. **26**(11): p. 4442-9.
88. Ferapontova, E.E., E.M. Olsen, and K.V. Gothelf, *An RNA aptamer-based electrochemical biosensor for detection of theophylline in serum.* J Am Chem Soc, 2008. **130**(13): p. 4256-8.
89. Stojanovic, M.N., P. de Prada, and D.W. Landry, *Aptamer-based folding fluorescent sensor for cocaine.* J Am Chem Soc, 2001. **123**(21): p. 4928-31.

90. Elowe, N.H., et al., *Small-molecule screening made simple for a difficult target with a signaling nucleic acid aptamer that reports on deaminase activity.* Angew Chem Int Ed Engl, 2006. **45**(34): p. 5648-52.
91. Luo, Y., et al., *Coordination polymer nanobelts for nucleic acid detection.* Nanotechnology, 2011. **22**(19): p. 195502.
92. Chang, H., et al., *Graphene fluorescence resonance energy transfer aptasensor for the thrombin detection.* Anal Chem, 2010. **82**(6): p. 2341-6.
93. Zhao, W., et al., *Simple and rapid colorimetric biosensors based on DNA aptamer and noncrosslinking gold nanoparticle aggregation.* Chembiochem, 2007. **8**(7): p. 727-31.
94. Blake, D.A., et al., *Antibody-based sensors for heavy metal ions.* Biosens Bioelectron, 2001. **16**(9-12): p. 799-809.
95. Bini, A., et al., *Analytical performances of aptamer-based sensing for thrombin detection.* Anal Chem, 2007. **79**(7): p. 3016-9.
96. Tombelli, S., et al., *Aptamer-based biosensors for the detection of HIV-1 Tat protein.* Bioelectrochemistry, 2005. **67**(2): p. 135-41.
97. Hwang, K.S., et al., *Nanomechanical microcantilever operated in vibration modes with use of RNA aptamer as receptor molecules for label-free detection of HCV helicase.* Biosens Bioelectron, 2007. **23**(4): p. 459-65.
98. Cella, L.N., et al., *Nano aptasensor for protective antigen toxin of anthrax.* Anal Chem, 2010. **82**(5): p. 2042-7.
99. Wang, L., et al., *Label-free, regenerative and sensitive surface plasmon resonance and electrochemical aptasensors based on graphene.* Chem Commun (Camb), 2011. **47**(27): p. 7794-6.
100. Vivekananda, J. and J.L. Kiel, *Anti-Francisella tularensis DNA aptamers detect tularemia antigen from different subspecies by Aptamer-Linked Immobilized Sorbent Assay.* Lab Invest, 2006. **86**(6): p. 610-8.
101. Xu, H., et al., *Aptamer-functionalized gold nanoparticles as probes in a dry-reagent strip biosensor for protein analysis.* Anal Chem, 2009. **81**(2): p. 669-75.
102. Crawley, J.T., et al., *The central role of thrombin in hemostasis.* J Thromb Haemost, 2007. **5 Suppl 1**: p. 95-101.
103. Fenton, J.W., 2nd, *Thrombin specificity.* Ann N Y Acad Sci, 1981. **370**: p. 468-95.
104. Shuman, M.A., *Thrombin-cellular interactions.* Ann N Y Acad Sci, 1986. **485**: p. 228-39.
105. Bode, W., *Structure and interaction modes of thrombin.* Blood Cells Mol Dis, 2006. **36**(2): p. 122-30.
106. Russo Krauss, I., et al., *Thrombin-aptamer recognition: a revealed ambiguity.* Nucleic Acids Res, 2011. **39**(17): p. 7858-67.
107. Shuman, M.A. and P.W. Majerus, *The measurement of thrombin in clotting blood by radioimmunoassay.* J Clin Invest, 1976. **58**(5): p. 1249-58.
108. Becker, R.C. and F.A. Spencer, *Thrombin: Structure, Biochemistry, Measurement, and Status in Clinical Medicine.* J Thromb Thrombolysis, 1998. **5**(3): p. 215-229.
109. Suo, Z., B.A. Citron, and B.W. Festoff, *Thrombin: a potential proinflammatory mediator in neurotrauma and neurodegenerative disorders.* Curr Drug Targets Inflamm Allergy, 2004. **3**(1): p. 105-14.
110. Turgeon, V.L. and L.J. Houenou, *The role of thrombin-like (serine) proteases in the development, plasticity and pathology of the nervous system.* Brain Res Brain Res Rev, 1997. **25**(1): p. 85-95.
111. Paborsky, L.R., et al., *The single-stranded DNA aptamer-binding site of human thrombin.* J Biol Chem, 1993. **268**(28): p. 20808-11.
112. Tasset, D.M., M.F. Kubik, and W. Steiner, *Oligonucleotide inhibitors of human thrombin that bind distinct epitopes.* J Mol Biol, 1997. **272**(5): p. 688-98.
113. Bock, L.C., et al., *Selection of single-stranded DNA molecules that bind and inhibit human thrombin.* Nature, 1992. **355**(6360): p. 564-6.

114. Macaya, R.F., et al., *Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution.* Proc Natl Acad Sci U S A, 1993. **90**(8): p. 3745-9.
115. Padmanabhan, K., et al., *The structure of alpha-thrombin inhibited by a 15-mer single-stranded DNA aptamer.* J Biol Chem, 1993. **268**(24): p. 17651-4.
116. Gatto, B., M. Palumbo, and C. Sissi, *Nucleic acid aptamers based on the G-quadruplex structure: therapeutic and diagnostic potential.* Curr Med Chem, 2009. **16**(10): p. 1248-65.
117. Baldrich, E. and C.K. O'Sullivan, *Ability of thrombin to act as molecular chaperone, inducing formation of quadruplex structure of thrombin-binding aptamer.* Anal Biochem, 2005. **341**(1): p. 194-7.
118. Ikebukuro, K., C. Kiyohara, and K. Sode, *Novel electrochemical sensor system for protein using the aptamers in sandwich manner.* Biosens Bioelectron, 2005. **20**(10): p. 2168-72.
119. Centi, S., et al., *Aptamer-based detection of plasma proteins by an electrochemical assay coupled to magnetic beads.* Anal Chem, 2007. **79**(4): p. 1466-73.
120. Tang, Q., X. Su, and K.P. Loh, *Surface plasmon resonance spectroscopy study of interfacial binding of thrombin to antithrombin DNA aptamers.* J Colloid Interface Sci, 2007. **315**(1): p. 99-106.
121. Lao, Y.H., K. Peck, and L.C. Chen, *Enhancement of aptamer microarray sensitivity through spacer optimization and avidity effect.* Anal Chem, 2009. **81**(5): p. 1747-54.
122. Zhao, Q., et al., *Aptamer-linked assay for thrombin using gold nanoparticle amplification and inductively coupled plasma-mass spectrometry detection.* Anal Chem, 2009. **81**(17): p. 7484-9.
123. Niu, S., et al., *Fluorescence detection of thrombin using autocatalytic strand displacement cycle reaction and a dual-aptamer DNA sandwich assay.* Anal Biochem, 2012. **421**(2): p. 362-7.
124. Cho, E.J., J.W. Lee, and A.D. Ellington, *Applications of aptamers as sensors.* Annu Rev Anal Chem (Palo Alto Calif), 2009. **2**: p. 241-64.
125. Wang, L., et al., *Simultaneously fluorescence detecting thrombin and lysozyme based on magnetic nanoparticle condensation.* Talanta, 2009. **79**(3): p. 557-61.
126. Marson, G., M. Palumbo, and C. Sissi, *Folding versus charge: understanding selective target recognition by the thrombin aptamers.* Curr Pharm Des, 2012. **18**(14): p. 2027-35.
127. Folkman, J., *Seminars in Medicine of the Beth Israel Hospital, Boston. Clinical applications of research on angiogenesis.* N Engl J Med, 1995. **333**(26): p. 1757-63.
128. Scappaticci, F.A., *Mechanisms and future directions for angiogenesis-based cancer therapies.* J Clin Oncol, 2002. **20**(18): p. 3906-27.
129. Bressler, S.B., *Introduction: Understanding the role of angiogenesis and antiangiogenic agents in age-related macular degeneration.* Ophthalmology, 2009. **116**(10 Suppl): p. S1-7.
130. Koch, A.E., *Angiogenesis as a target in rheumatoid arthritis.* Ann Rheum Dis, 2003. **62 Suppl 2**: p. ii60-7.
131. Gatto, B. and M. Cavalli, *From proteins to nucleic acid-based drugs: the role of biotech in anti-VEGF therapy.* Anticancer Agents Med Chem, 2006. **6**(4): p. 287-301.
132. Wiesmann, C., et al., *Crystal structure of the complex between VEGF and a receptor-blocking peptide.* Biochemistry, 1998. **37**(51): p. 17765-72.
133. Wiesmann, C., et al., *Crystal structure at 1.7 A resolution of VEGF in complex with domain 2 of the Flt-1 receptor.* Cell, 1997. **91**(5): p. 695-704.
134. Muller, Y.A., et al., *The crystal structure of vascular endothelial growth factor (VEGF) refined to 1.93 A resolution: multiple copy flexibility and receptor binding.* Structure, 1997. **5**(10): p. 1325-38.
135. Fairbrother, W.J., et al., *1H, 13C, and 15N backbone assignment and secondary structure of the receptor-binding domain of vascular endothelial growth factor.* Protein Sci, 1997. **6**(10): p. 2250-60.
136. Fairbrother, W.J., et al., *Solution structure of the heparin-binding domain of vascular endothelial growth factor.* Structure, 1998. **6**(5): p. 637-48.

137. Pan, B., et al., *Solution structure of a phage-derived peptide antagonist in complex with vascular endothelial growth factor.* J Mol Biol, 2002. **316**(3): p. 769-87.
138. Ferrara, N., H.P. Gerber, and J. LeCouter, *The biology of VEGF and its receptors.* Nat Med, 2003. **9**(6): p. 669-76.
139. Bhisitkul, R.B., *Vascular endothelial growth factor biology: clinical implications for ocular treatments.* Br J Ophthalmol, 2006. **90**(12): p. 1542-7.
140. Hasegawa, H., K. Sode, and K. Ikebukuro, *Selection of DNA aptamers against VEGF165 using a protein competitor and the aptamer blotting method.* Biotechnol Lett, 2008. **30**(5): p. 829-34.
141. Nonaka, Y., K. Sode, and K. Ikebukuro, *Screening and improvement of an anti-VEGF DNA aptamer.* Molecules, 2010. **15**(1): p. 215-25.
142. Sissi, C., B. Gatto, and M. Palumbo, *The evolving world of protein-G-quadruplex recognition: a medicinal chemist's perspective.* Biochimie, 2011. **93**(8): p. 1219-30.
143. Kwon, O.S., S.J. Park, and J. Jang, *A high-performance VEGF aptamer functionalized polypyrrole nanotube biosensor.* Biomaterials, 2010. **31**(17): p. 4740-7.
144. Cruz-Aguado, J.A. and G. Penner, *Fluorescence polarization based displacement assay for the determination of small molecules with aptamers.* Anal Chem, 2008. **80**(22): p. 8853-5.
145. Sosic, A., et al., *Human thrombin detection through a sandwich aptamer microarray: interaction analysis in solution and in solid phase.* Sensors (Basel), 2011. **11**(10): p. 9426-41.
146. Neidle, S., *The structures of quadruplex nucleic acids and their drug complexes.* Curr Opin Struct Biol, 2009. **19**(3): p. 239-50.
147. Baldrich, E., A. Restrepo, and C.K. O'Sullivan, *Aptasensor development: elucidation of critical parameters for optimal aptamer performance.* Anal Chem, 2004. **76**(23): p. 7053-63.
148. Ponikova, S., M. Antalik, and T. Hianik, *A circular dichroism study of the stability of guanine quadruplexes of thrombin DNA aptamers at presence of K+ and Na+ ions.* Gen Physiol Biophys, 2008. **27**(4): p. 271-7.
149. Fialova, M., J. Kypr, and M. Vorlickova, *The thrombin binding aptamer GGTTGGTGTGGTTGG forms a bimolecular guanine tetraplex.* Biochem Biophys Res Commun, 2006. **344**(1): p. 50-4.
150. Sahni, A. and C.W. Francis, *Vascular endothelial growth factor binds to fibrinogen and fibrin and stimulates endothelial cell proliferation.* Blood, 2000. **96**(12): p. 3772-8.
151. Keyt, B.A., et al., *The carboxyl-terminal domain (111-165) of vascular endothelial growth factor is critical for its mitogenic potency.* J Biol Chem, 1996. **271**(13): p. 7788-95.
152. Hornbeck, P., *Enzyme-linked immunosorbent assays.* Curr Protoc Immunol, 2001. **Chapter 2**: p. Unit 2 1.
153. Paramasivan, S., I. Rujan, and P.H. Bolton, *Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding.* Methods, 2007. **43**(4): p. 324-31.
154. Huang, D.W., et al., *Time-resolved fluorescence aptamer-based sandwich assay for thrombin detection.* Talanta. **83**(1): p. 185-9.
155. Tennico, Y.H., et al., *On-chip aptamer-based sandwich assay for thrombin detection employing magnetic beads and quantum dots.* Anal Chem. **82**(13): p. 5591-7.
156. Zeng, X., Z. Shen, and R. Mernaugh, *Recombinant antibodies and their use in biosensors.* Anal Bioanal Chem, 2012. **402**(10): p. 3027-38.
157. Huang, D.W., et al., *Time-resolved fluorescence aptamer-based sandwich assay for thrombin detection.* Talanta, 2010. **83**(1): p. 185-9.
158. Fang, L., et al., *Quantitative electrochemiluminescence detection of proteins: Avidin-based sensor and tris(2,2'-bipyridine) ruthenium(II) label.* Biosens Bioelectron, 2008. **23**(11): p. 1645-51.
159. Chen, Y., et al., *On-chip synthesis of RNA aptamer microarrays for multiplexed protein biosensing with SPR imaging measurements.* Langmuir, 2012. **28**(22): p. 8281-5.

APPENDIX

(12-55) NC peptide: structure and biological activity

The nucleocapsid protein NCp7 of human immunodeficiency virus type 1 (HIV-1) is a small, basic protein that is tightly associated with genomic RNA in the mature infectious virus. NCp7 is characterized by two copies of a highly conserved retroviral-type, zinc-finger motif, also called CCHC motif that binds zinc ions strongly in a constrained structure. In order to examine the nature and the extent of NC destabilizing activity, we investigated by a fluorescence assay the interaction of cTAR with the (12-55)NC peptide containing the zinc-finger motifs but lacking the ability to aggregate the oligonucleotides; further more we characterized the structure of the peptide using circular dichroic spectroscopy in both the presence and absence of zinc ions.

Analysis of the structure

The peptide synthesis was very difficult for the provider. We received different stocks of peptide: an aliquot of peptide (12-55)NC was synthesized in the absence of zinc ions; another aliquot of peptide was synthesized in the presence of $ZnCl_2$ to preserve the highly oxidizable Cys residues and the peptide, named (12-55)NCZn was lyophilized in the zinc-bound form. The lyophilized peptides (12-55)NC and (12-55)NCZn were resuspended in H_2O or in TN 10/20 (Tris-HCl 10mM, NaCl 20mM pH 7.5).

The early spectra analysis was obtained using a Nanodrop spectrophotometer (Figure I).

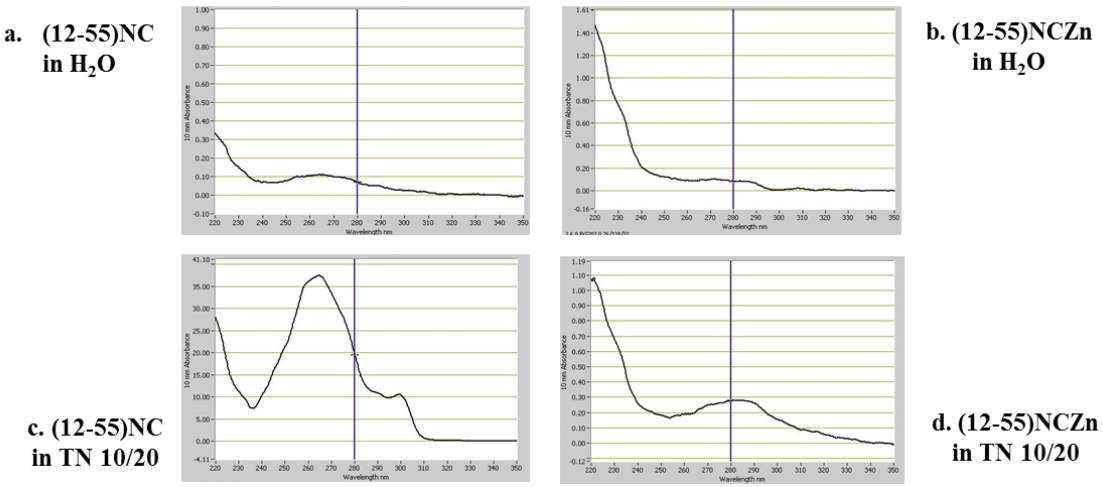


Figure I UV-VIS spectra obtained using a Nanodrop spectrophotometer. a) (12-55)NC peptide resuspended in H_2O; b) (12-55)NCZn peptide resuspended in H_2O; c) (12-55)NC peptide resuspended in TN 10/20; d) (12-55)NCZn peptide resuspended in TN 10/20. Peptide concentration was determined using an extinction coefficient at 280nm of 5700 M^{-1} cm^{-1}.

Then, peptide concentrations obtained by Nanodrop spectrophotometer, were confirmed in H_2O using a UV-Vis Spectrophotometer: the obtained spectra are shown in Figure II. Lyophilized peptide was not completely soluble after the first resuspension (blue line), hence the precipitate was resuspended and analyzed again (red line).

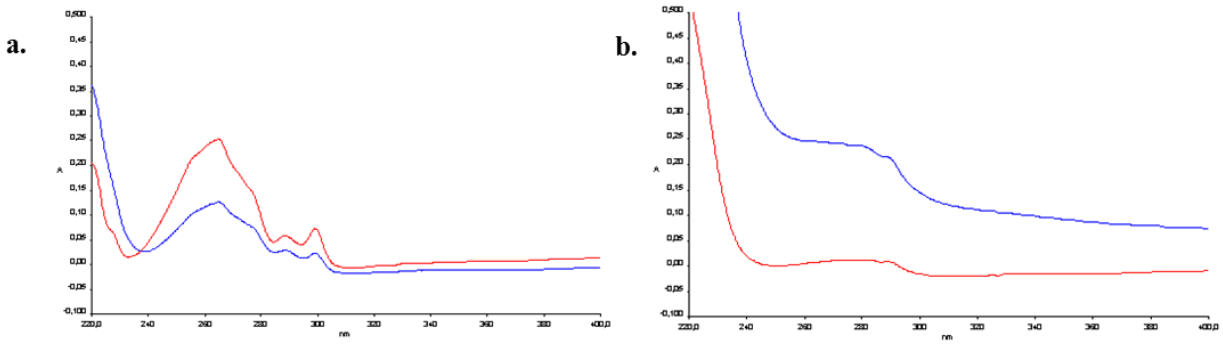


Figure II a) UV-Vis analysis of (12-55)NC peptide resuspended in H_2O; b) UV-Vis analysis of (12-55)NCZn peptide resuspended in H_2O.

(12-55)NC was dissolved two times to obtain a higher concentration of peptide; the most amount of (12-55)NCZn was instead obtained after the first resuspension. In both cases was detected an unpredicted absorbance at 260 nm: it can be associated to a contaminant species present in the peptide samples. This hypothesis was confirmed using LC-MS, Mariner Applied Biosystem (data not shown). The presence of an unknown contaminant in the samples prevented the correct valuation of the peptide concentration.

The correct folding of the peptide is indispensable for the chaperone activity of NC. For this reason, the analysis of the secondary structure of the peptide by circular dichroic spectroscopy (CD) was performed prior to investigate the biological activity of both (12-55)NC and (12-55)NCZn. In Figure III are reported the dichroic spectra obtained by the analysis of the (12-55)NC peptide (Figure IIIa) and the (12-55)NCZn peptide (Figure IIIb). The exactly concentrations of peptides were not determined due to the presence of an unknown contaminant in the samples and for this reason these spectra were not normalized for the peptide concentration. In the case of NC, with or without zinc ions, the signal was not significant because the amount of peptide was too low. In the case of NCZn, which was treated also with EDTA, we can observe a different fold from that observed for NC without zinc ions.

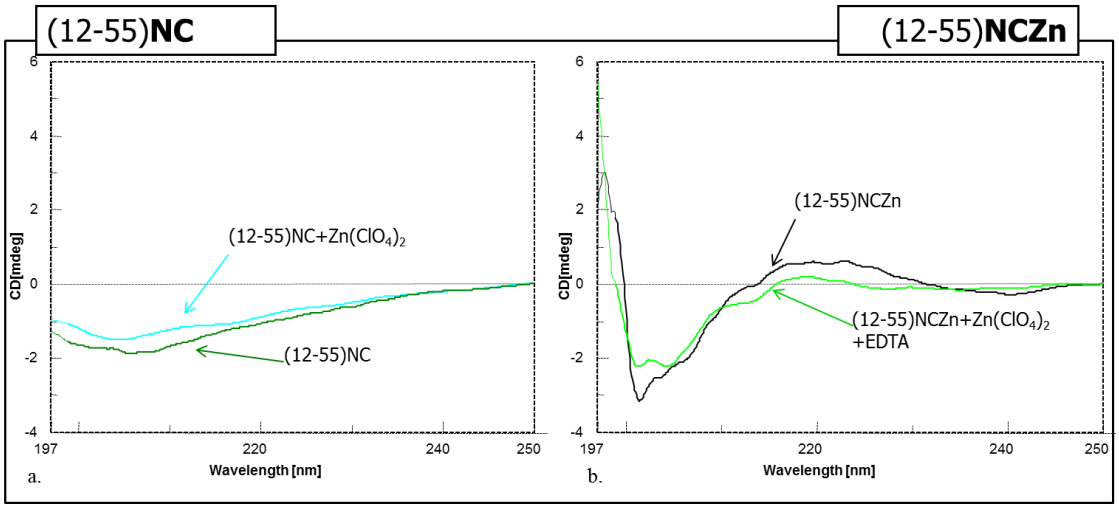


Figure III Circular dichroism (CD) spectra of the (12-55)NC (a) and of the (12-55)NCZn (b), in the presence and in the absence of zinc ions.

Analysis of the biological activity

Even if the exactly concentration of peptide was not determined due to the presence of an unknown contaminant in the samples, a fluorescence assay was performed in order to evaluate the biological activity of the synthesized (12-55)NC and (12-55)NCZn peptides. The doubly labeled cTAR (5'-FAM and 3'-DAB) was used. The fluorescence signal of the folded cTAR was analyzed in cuvette with 494 nm as excitation and 500-700 nm as emission wavelengths. Increasing concentrations of (12-55)NCZn, resuspended in H_2O, were added to cTAR before reading. The fluorescence spectra are reported in Figure IV.

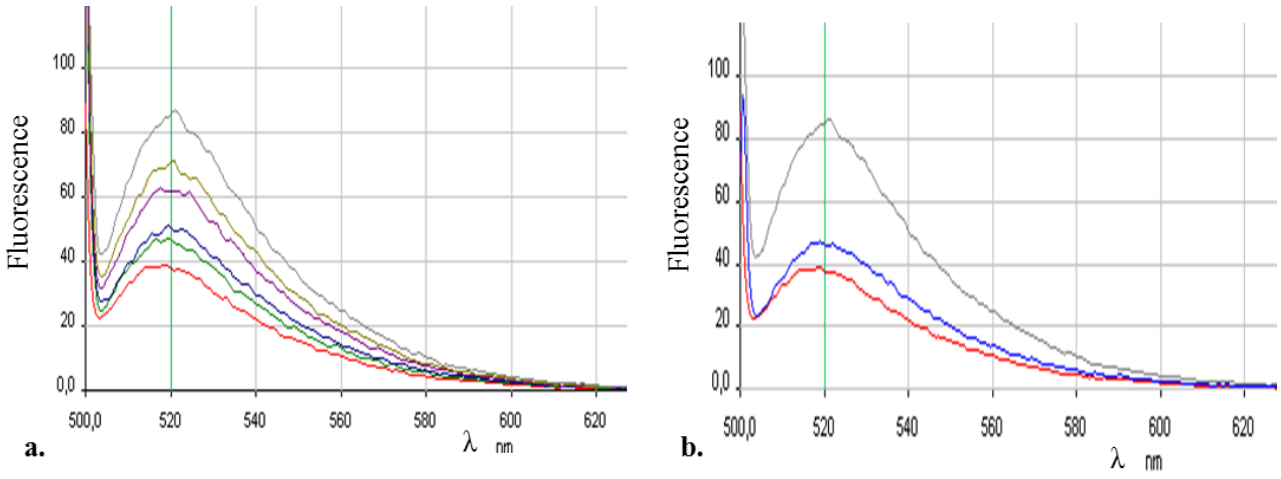


Figure IV Fluorescence Assay: analysis of the biological activity of (12-55)NCZn. a) Fluorescence spectra of 0.1μM 3'-FAM and 5'-DAB modified cTAR in TNMg (Tris 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM, pH 7.5) in absence (red line) and in presence (from green to gray lines) of increasing concentrations of (12-55)NCZn. Spectra were corrected for buffer fluorescence. **b)** Fluorescence spectra of 0.1μM 3'-FAM and 5'-DAB modified cTAR in TNMg (Tris 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM, pH 7.5) in absence of (12-55)NCZn (red line), in presence of (12-55)NCZn (gray line) and after adding an excess of EDTA (blu line), which removes Zn ions from Zn-fingers and inactivate the peptide. Spectra were corrected for buffer fluorescence.

We can observe that increasing concentrations of (12-55)NC determines a fluorescence increase. The fluorescence increase in the presence of (12-55)NC is due to the increase of the distance between DAB and FAM that accompanies the destabilization of the cTAR stem-structure. Therefore the (12-55)NCZn peptide can act as chaperone in these conditions.
In Figure IVb, we can observe the effect of the addition of EDTA. The addiction of EDTA determines a fluorescence decrease. EDTA is a chelating agent that can remove Zn ions from the Zn-finger motif and inactivate the peptide. If the peptide is not correct folded, it cannot act as chaperone on cTAR.

Dialysis of (12-55)NCZn

The first stocks of (12-55)NC and (12-55)NCZn was dialyzed to remove the contaminant from the peptide samples using a dialysis cassette (Slide-A-Lyzer Dialysis Cassettes, ThermoScientific). The used dialysis solution was in both cases $Zn(ClO_4)_2$ 600 µM. After dialysis the samples were dried and then resuspended in H_2O.
Finally, the biological activity of the dialyzed (12-55)NCZn was investigated by the fluorescence assay on cTAR (Figure Va).

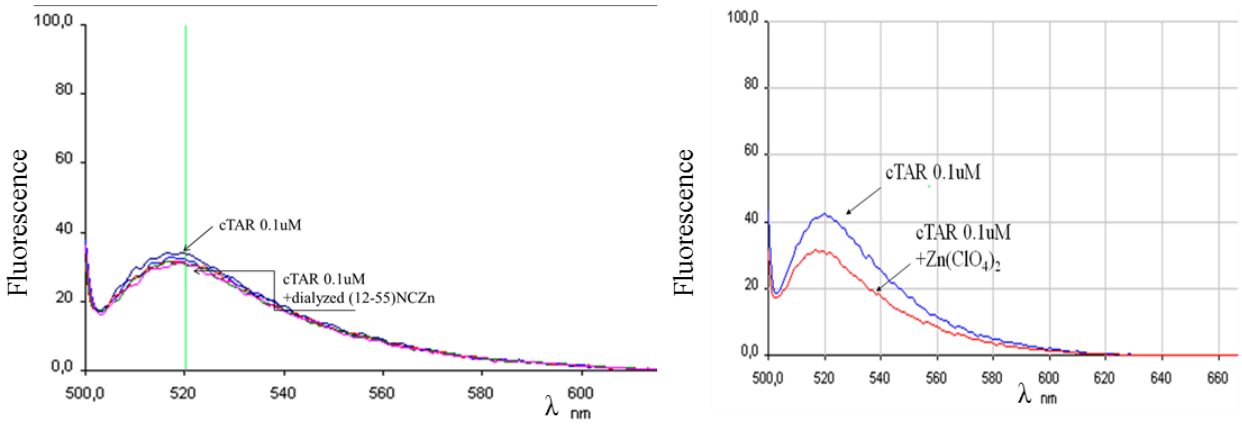


Figure V Fluorescence Assay. a) Analysis of the biological activity of the dialyzed (12-55)NCZn, first stock. Fluorescence spectra of 0.1µM 3'-FAM and 5'-DAB modified cTAR in TNMg (Tris 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM, pH 7.5) in TNMg in absence of NCZn (blu line) and in presence of increasing concentration of the dyalized NCZn (pink line). **b) Analysis of the effect of $Zn(ClO_4)_2$ on cTAR fluorescence.** Fluorescence spectra of 0.1µM 3'-FAM and 5'-DAB modified cTAR, in absence of NC, in TNMg (Tris 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM, pH 7.5) (blu line) and cTAR, in absence of NC, in TNMg (Tris 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM, pH 7.5) in presence of 1mM $Zn(ClO_4)_2$ (red line).

The addiction of the dialyzed peptide determines a fluorescence quench. This result is not according with the fluorescence assay performed using the not dialyzed stock of (12-55)NCZn (Figure IVa). The only difference between the two experiments is the presence of an excess of $Zn(ClO_4)_2$ in the second experience: the first stock of NCZn was resuspended in

water and analyzed itself; the second analysis was instead performed using the same stock of peptide, resuspended in water and dialyzed towards $Zn(ClO_4)_2$.
In Figure Vb is reported the analysis of the effect of $Zn(ClO_4)_2$ on cTAR fluorescence: we can observe that the addiction of Zinc perclorate to cTAR determines a significant fluorescence decrease.

We analyzed also the free fluorescein to confirm the fluorescence quenching by $Zn(ClO_4)_2$: in presence of Zinc perclorate the fluorescence of the fluorescein decreased (Figure VI). Therefore, we can affirm that the divalent cation Zn^{2+} is responsible of the fluorescence quench.

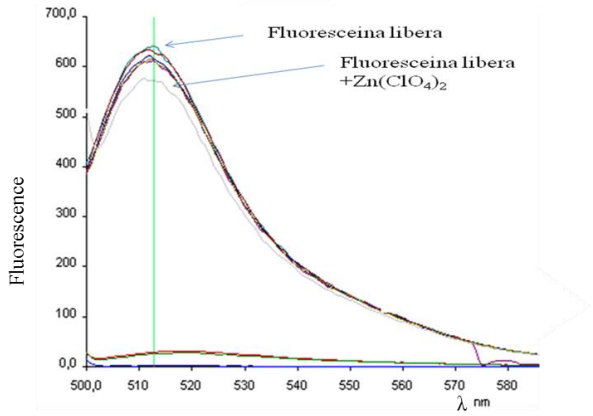


Figure VI Fluorescence Assay: analysis of the effect of $Zn(ClO_4)_2$ on free fluorescein. Fluorescence spectra of 0.1µM fluorescein, in TNMg (Tris 10 mM, NaCl 20 mM, $Mg(ClO_4)_2$ 1 mM, pH 7.5) in absence of Zn ions and in presence of increasing concentration of $Zn(ClO_4)_2$.

Therefore, we can affirm that the divalent cation Zn^{2+} in solution is responsible of the fluorescence quench. Literature data supported our hypothesis (Rupcich et al. 2006).